Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **419,** 2797–2820 (2012) doi:10.1111/j.1365-2966.2011.19931.x

The evolution of active galactic nuclei across cosmic time: what is downsizing?

N. Fanidakis,[1]⋆ C. M. Baugh,[1] A. J. Benson,[2] R. G. Bower,[1] S. Cole,[1] C. Done,[1] C. S. Frenk,[1] R. C. Hickox,[1] C. Lacey[1] and C. del P. Lagos[1]

[1]*Institute for Computational Cosmology, Department of Physics, University of Durham, Science Laboratories, South Road, Durham DH1 3LE*
[2]*Theoretical Astrophysics, California Institute of Technology, Mail Code 130-33, 1200 East California Boulevard, Pasadena, CA 91125, USA*

Accepted 2011 September 28. Received 2011 July 13; in original form 2010 November 22

ABSTRACT

We use a coupled model of the formation and evolution of galaxies and black holes (BHs) to study the evolution of active galactic nuclei (AGNs) in a cold dark matter universe. The model is embedded in the galaxy formation code GALFORM and predicts the masses, spins and mass accretion histories of BHs in tandem with the formation of their host galaxies. BHs grow by accretion during starbursts, triggered by discs becoming dynamically unstable or by galaxy mergers, and accretion from quasi-hydrostatic hot gas haloes. Using an empirical law for AGN obscuration, our model matches the observed luminosity functions (LFs) of AGNs over a wide range of redshifts. Due to the suppression of cooling in massive haloes by AGN feedback, at low redshift ($z \lesssim 2$), the brightest quasars ($L_{\rm bol} \gtrsim 10^{46}\,{\rm erg\,s^{-1}}$) are predicted preferentially to inhabit haloes with masses $\simeq 10^{12}$–$10^{13}\,{\rm M_\odot}$. The model predicts a hierarchical buildup of BH mass, with the typical mass of actively growing BHs increasing with decreasing redshift. Nevertheless, the model displays clear 'downsizing' as reflected in the differential evolution of the space density of faint and bright AGNs. This arises naturally from the interplay between the starburst and hot gas halo accretion modes. The faint end of the LF is dominated by massive BHs accreting at low rates via a thick disc, primarily during the hot-halo mode. The bright end is populated by BHs accreting close to or above the Eddington limit during the starburst mode. Obscuration plays a central role in determining the observed abundance of AGNs and, hence, in their implied cosmic evolution.

Key words: methods: numerical – galaxies: active – galaxies: nuclei – quasars: general.

1 INTRODUCTION

The evolution of active galactic nuclei (AGNs) provides a crucial test of galaxy formation models. Not only are AGNs an important component of the Universe, but they are also thought to play a key role in the evolution of their host galaxies and their environment. For example, feedback processes which accompany the growth of black holes (BHs) during AGN activity are thought to influence star formation (SF) through the suppression of cooling flows in galaxy groups and clusters (Dalla Vecchia et al. 2004; Springel, Di Matteo & Hernquist 2005b; Bower et al. 2006; Croton et al. 2006; Hopkins et al. 2006; Thacker, Scannapieco & Couchman 2006; Lagos, Cora & Padilla 2008; Somerville et al. 2008; see also Marulli et al. 2008). Similarly, energy feedback from AGNs is a key ingredient in determining the X-ray properties of the intracluster medium (Bower, McCarthy & Benson 2008; McCarthy et al. 2010). Understanding these processes requires that the underlying galaxy formation model be able to track and reproduce the evolution of AGNs.

Early surveys demonstrated that quasi-stellar objects (QSOs) undergo significant evolution from $z \sim 0$ up to $z \sim 2$ (Schmidt & Green 1983; Boyle, Shanks & Peterson 1988; Hewett, Foltz & Chaffee 1993; Boyle et al. 2000). Beyond $z \sim 2$, the space density of QSOs starts to decline (Warren, Hewett & Osmer 1994; Schmidt, Schneider & Gunn 1995; Fan et al. 2001a; Wolf et al. 2003). Recently, there has been much progress in pinning down the evolution of faint AGNs in the optical. Using the 2-degree-field (2dF) QSO Redshift survey (2QZ), the luminosity function (LF) was probed to around a magnitude fainter than the break to $z \sim 2$ (Croom et al. 2001, 2004). This limit was extended by the 2dF-SDSS (Sloan Digital Sky Survey) luminous red galaxy and QSO (2SLAQ) survey (Richards et al. 2005; Croom et al. 2009a). The estimate of the LF from the final 2SLAQ catalogue reached $M_{b_J} \simeq -19.8$ at $z = 0.4$ (Croom et al. 2009b). With the 2SLAQ LF, Croom et al. (2009b)

⋆E-mail: nikolaos.fanidakis@dur.ac.uk

© 2011 The Authors
Monthly Notices of the Royal Astronomical Society © 2011 RAS

demonstrated that in the optical, faint quasars undergo mild evolution, with their number density peaking at lower redshifts than is the case for bright quasars (see also Bongiorno et al. 2007).

Faint AGNs can be selected robustly in X-rays (Hasinger et al. 2001; Giacconi et al. 2002; Alexander et al. 2003; Barcons et al. 2007, see also the review by Brandt & Hasinger 2005), which means that a wider range of AGN luminosity can be probed in X-rays than is possible in the optical. This permits the study of the evolution of a wider variety of AGNs in addition to QSOs (e.g. Seyfert galaxies) and thus provides a more representative picture of the various AGN populations. The evolution of the AGN LF in X-rays has been investigated by many authors by employing data from the *HEAO-1*, *ASCA*, *ROSAT*, *Chandra* and *XMM–Newton* surveys (Miyaji, Hasinger & Schmidt 2000; La Franca et al. 2002; Cowie et al. 2003; Fiore et al. 2003; Ueda et al. 2003; Barger et al. 2005; Hasinger, Miyaji & Schmidt 2005). These studies show that in both soft (0.5–2 keV) and hard (2–10 keV) X-rays, faint AGNs are found to evolve modestly with redshift. In contrast, bright AGNs show strong evolution, similar to that seen for quasars in the optical. In addition, observations in soft X-rays (SXs) suggest that the comoving space density of bright AGNs peaks at higher redshifts ($z \sim 2$) than faint AGNs ($z < 1$; Hasinger et al. 2005).

The differential evolution of bright and faint AGNs with redshift has been described as *downsizing* (Barger et al. 2005; Hasinger et al. 2005). This implies that AGN activity in the low-z Universe is dominated by either high-mass BHs accreting at low rates or low-mass BHs growing rapidly. Hopkins et al. (2005b) proposed that the faint end of the LF is composed of high-mass BHs experiencing quiescent accretion (see also Hopkins et al. 2005a,c; Babić et al. 2007). The bright end, on the other hand, corresponds in this picture to BHs accreting near their Eddington limit. In the Hopkins et al. model, quasar activity is short-lived and is assumed to be driven by galaxy mergers (Di Matteo, Springel & Hernquist 2005).

The masses of accreting BHs can be estimated using the spectra of the AGNs. Quasars are ideal for this since they are detected up to very high redshifts and hence can trace the evolution of actively growing BHs back into the early Universe (Fan et al. 2001b, 2003; Fontanot et al. 2007; Willott et al. 2010). The BH masses are derived from optical or ultraviolet (UV) spectroscopy using empirical relations. In particular, mass-scaling relations between the widths of different broad emission lines and continuum luminosities that have been calibrated against reverberation mapping results (Vestergaard 2002; Vestergaard & Peterson 2006) have allowed BH mass estimates in several large, unobscured (type-1) quasar samples. When translating quasar luminosities into BH masses using the width of broad lines, a similar downsizing is seen in BH mass (Vestergaard & Osmer 2009; Kelly et al. 2010), suggesting that the most massive BHs ($M_{\rm BH} > 10^9\,{\rm M}_\odot$) were already in place at $z > 2$, whereas the growth of the less massive ones is delayed to lower redshifts.

An important feature of AGNs which should be taken into account is that they exhibit evidence of obscuration at both optical and SX wavelengths. The obscuration may be linked to the existence of a geometrical torus around the accretion disc whose presence is invoked in the AGN unification scheme (Antonucci 1993; Urry & Padovani 1995) or to intervening dust clouds related to physical processes within the host galaxy (Martínez-Sansigre et al. 2005; Ballantyne, Everett & Murray 2006b, but also Goulding & Alexander 2009; Lagos et al. 2011b). As a consequence, a large fraction of AGNs could be obscured and thus missing from optical and SX surveys. Hence, when applying the scaling relations to estimate the masses of accreting BHs, the absence of obscured (type-2) quasars from the AGN samples may introduce significant biases into the inferred evolution of BH mass. Only hard X-rays (HXs) can directly probe the central engine by penetrating the obscuring medium, therefore providing complete and unbiased samples of AGNs (Ueda et al. 2003; Barger et al. 2005; La Franca et al. 2005). However, even in HXs, a population of Compton-thick sources, namely AGNs with hydrogen column densities exceeding $N_{\rm H} \simeq 10^{24}\,{\rm cm}^{-2}$, would still be missing (Comastri 2004; Alexander et al. 2005, 2008; Goulding et al. 2011).

In the past decade, many semi-analytical models have attempted to explain the properties and evolution of AGNs. These models incorporate simple schemes for the growth of supermassive BHs (SMBHs) and assume that AGN activity is suavely triggered by major galaxy mergers or disc instabilities (Kauffmann & Haehnelt 2000; De Lucia & Blaizot 2007; Malbon et al. 2007; Monaco, Fontanot & Taffoni 2007; see also Granato et al. 2004; Lapi et al. 2006 for an alternative approach). An important feature of these models is the inclusion of AGN feedback, a physical mechanism necessary for reproducing the morphology and colour of massive galaxies, and also the steep cut-off of the galaxy LF at low redshift (Bower et al. 2006; Croton et al. 2006; Lagos et al. 2008). These models can provide predictions for the statistical properties of AGNs, such as the LF (Marulli et al. 2008; Croton 2009) and correlation function (Bonoli et al. 2009, 2010), by making simple assumptions for the properties of accretion on to the BHs (Shankar et al. 2004; Hopkins et al. 2005a).

In this paper, we present a study of AGN evolution using semi-analytic modelling (see Baugh 2006 for a review). Our aim is to provide a robust framework for understanding the downsizing of AGNs within a self-consistent galaxy formation model. This paper is organized as follows. In Section 2, we present the galaxy formation model upon which we build our AGN model. In Section 3, we study the evolution of BH mass and explore the scaling relations predicted between the BH mass and the mass of the host galaxy and dark matter (DM) halo. In Section 4, we present the essential ingredients of the AGN model and study the evolution of the physical properties which, together with the mass, provide a complete description of accreting BHs. In Section 5, we present our predictions for the evolution of the optical, SX/HX and bolometric LFs. Finally, in Section 6, we explore the downsizing of AGNs in our model. The cosmology adopted in the simulations is $\Omega_{\rm m} = 0.25$, $\Omega_{\rm b} = 0.045$, $\Omega_\Lambda = 0.75$ and $\sigma_8 = 0.9$.[1] Throughout this paper, we set $h = 0.7$ in the galaxy properties we calculate.

2 THE GALAXY FORMATION MODEL

We use the GALFORM semi-analytical galaxy formation code (Cole et al. 2000, also Baugh et al. 2005; Bower et al. 2006) and the extension to follow the evolution of BH mass and spin introduced by Fanidakis et al. (2011). GALFORM simulates the formation and evolution of galaxies and BHs in a hierarchical cosmology by modelling a wide range of physical processes, including gas cooling, AGN heating, SF, supernova (SN) feedback, chemical evolution, and galaxy mergers. Apart from some minor updates in the modelling of the different physical processes, which are described throughout this section, the galaxy formation part (and associated parameters) of the model is kept as in Bower et al.

[1] $\Omega_{\rm m}$, $\Omega_{\rm b}$ and Ω_Λ express the present density of the baryonic, total matter and dark energy components of the Universe, respectively, relative to the critical density ($\rho_{\rm crit} = 3H^2/8\pi G$). σ_8 measures the rms mass fluctuations in spheres of radius $8\,h^{-1}$Mpc linearly extrapolated to the present day.

Our starting point is the Bower et al. (2006) galaxy formation model. This model invokes AGN feedback to suppress the cooling of gas in DM haloes with quasi-static hot atmospheres and has been shown to reproduce many observables, such as galaxy colours, stellar masses and LFs remarkably well. The model adopts a BH growth recipe based on that introduced by Malbon et al. (2007), and extended by Bower et al. (2006) and Fanidakis et al. (2011). During starbursts triggered by a disc instability (Efstathiou, Lake & Negroponte 1982) or galaxy merger, the BH accretes a fixed fraction, $f_{\rm BH}$, of the cold gas that is turned into stars in the burst, after taking into account SN feedback and recycling. The value of $f_{\rm BH}$ is chosen to fit the amplitude of the local observed $M_{\rm BH}$–$M_{\rm bulge}$ relation. In addition to the cold gas channel, quiescent accretion of gas from the hot halo during AGN feedback also contributes to the mass of the BH (see White & Frenk 1991; Cole et al. 2000; Croton et al. 2006, for the cooling properties of gas in haloes). To distinguish between these two accretion channels, we refer to the accretion triggered by disc instabilities or galaxy mergers as the *starburst mode* and the accretion from quasi-hydrostatic hot haloes as the *hot-halo mode*. These correspond to the *quasar* and *radio* modes, respectively, in the terminology used by Croton et al. (2006). Finally, mergers between BHs, which occur when the galaxies which host the BHs merge, redistribute BH mass and contribute to the buildup of the most massive BHs in the universe (see Fanidakis et al. 2011).

GALFORM calculates a plethora of properties for each galaxy including disc and bulge sizes, luminosities, colours and metallicities to list but a few. In this analysis, we are primarily interested in the properties that describe the BHs. Specifically, we output the BH mass, $M_{\rm BH}$, the amount of gas accreted in a starburst or during the hot-halo mode, $M_{\rm acc}$, the time that has passed since the last burst of SF experienced by the host galaxy, and the BH spin (discussed in more detail below). The time since the last burst is necessary to determine the beginning of the active phase of BH growth in the starburst mode.

Fanidakis et al. (2011) extended the model to track the evolution of BH spin, a. By predicting the spin, we can compute interesting properties such as the efficiency, ϵ, of converting matter into radiation during the accretion of gas on to a BH and the mechanical energy of jets in the Blandford & Znajek (1977) and Blandford & Payne (1982) mechanisms. The mechanical jet energy depends strongly on the accretion mode, as it most likely depends on the vertical (poloidal) magnetic field component close to the SMBH horizon. In the Fanidakis et al. (2011) model (hereinafter F11a model), we assume that the flow forms a geometrically thin disc for relatively high accretion rates (Shakura & Sunyaev 1973), switching at lower accretion rates to an advection-dominated accretion flow (ADAF: Narayan & Yi 1994; see Section 4.1 for further discussion). The collapse by two orders of magnitude in the scaleheight of the flow during the transition from an ADAF to a thin disc results in a similar drop in radio power. To calculate the evolution of the spin parameter, a, we use the *chaotic* accretion model (Berti & Volonteri 2008; King, Pringle & Hofmann 2008; Fanidakis et al. 2011). In this model, during the active phase gas is accreted on to the BH through a series of randomly oriented accretion discs, whose mass is limited by their self-gravity (King et al. 2005). We choose the chaotic accretion model [over the alternative, *prolonged* accretion model in the literature by Volonteri, Sikora & Lasota (2007) because the predicted BH spin distributions, through their influence on the strength of relativistic jets in the different accretion modes, reproduce better the population of radio-loud AGNs in the local Universe, Fanidakis et al. (2011)]. A correction to the mass of gas that is accreted is applied in order to account for the fraction of gas that turns into radiation during the accretion process. Finally, we note that we do not take into account BH ejection via gravitational-wave recoils during galaxy mergers (Merritt et al. 2004). This omission is not expected to have a significant impact on the evolution of BH mass in our model (Libeskind et al. 2006).

With this extension to the Bower et al. model, in Fanidakis et al., we were able to reproduce the diversity of nuclear activity seen in the local Universe. Furthermore, we demonstrated that the bulk of the phenomenology of AGNs can be naturally explained in a Λ cold dark matter (ΛCDM) universe by the coeval evolution of galaxies and BHs, coupled by AGN feedback. In this paper, we aim to extend the predictive power of the model to high redshifts and different wavelengths, to provide a complete and self-consistent framework for the formation and evolution of AGNs in a ΛCDM cosmology.

For completeness, we now list the model parameters which have an influence over the growth of BH mass:

(i) $f_{\rm BH}$, the fraction of the mass of stars produced in a starburst, after taking into account SN feedback and the recycling of gas, that is accreted on to the BH during the burst.

(ii) $f_{\rm Edd}$, the fraction of the Eddington luminosity of an accreting BH that is available to heat the hot halo during an episode of AGN feedback due to hot gas accretion.

(iii) $\epsilon_{\rm kin}$, the average kinetic efficiency of the jet during the hot-halo mode.

The fiducial model in this paper is denoted by F11b. We adopt the values for the above parameters that were used in the Bower et al. (2006) model, as listed in Table 1 (F11b model). Fanidakis et al. (2011) made some small changes to these parameter values, which are given in the F11a entry in Table 1. These changes resulted in a small change in the distribution of accretion rates in the hot-halo mode. In this analysis, we use the original parameter values in order to be consistent with the version of the Bower et al. model used by Lagos et al. (2011c) in their study of the effect of SF prescriptions on galaxy properties (see the following paragraph). By reverting to them from the values in F11a, there is a small tail of objects accreting in the hot-halo mode, for which the accretion rate is higher than that typically associated with accretion via a thick disc.

A further difference between our fiducial model, F11b, and the F11a model is the use of an improved SF law. Following Lagos et al. (2011c), who implemented a range of empirical and theoretical SF laws into GALFORM, we use the SF law of Blitz & Rosolowsky (2006, hereinafter BR06). Lagos et al. found that this SF law in particular improved the agreement between the model predictions and the observations for the mass function of atomic hydrogen in galaxies and the hydrogen mass-to-luminosity ratio (see also Lagos et al.

Table 1. Summary of the revised parameter values in the variants of the Bower et al. model considered here. Note that F11a uses a different model for quiescent SF in discs.

Study	$f_{\rm Edd}$ [a]	$\epsilon_{\rm kin}$ [b]	$f_{\rm BH}$ [c]	$\nu_{\rm SF}$ [d]
F11b (this study)	0.039	0.016	0.005	3.0
F11a (Fanidakis et al. 2011)	0.01	0.1	0.017	–
Lagos et al. (2011c)	0.039	0.016	0.005	5.25

[a] Fraction of the Eddington luminosity available for heating the gas in the host halo.
[b] Average kinetic efficiency of the jet during the hot-halo mode.
[c] Ratio of the mass of cold gas accreted on to the BH to the mass of cold gas used to form stars.
[d] SFR per unit molecular gas mass.

© 2011 The Authors, MNRAS **419,** 2797–2820
Monthly Notices of the Royal Astronomical Society © 2011 RAS

2011a for predictions on the evolution of the cold gas content of galaxies). The BR06 model has the attraction that it is more physical than the previous parametric SF law used in Bower et al., and agrees with observations of the relation between the surface density of SF and gas in galaxies. The BR06 law distinguishes between molecular and atomic hydrogen, with only the molecular hydrogen taking part in SF. The fraction of hydrogen in molecular form depends on the pressure within the galactic disc, which in turn is derived from the mass of gas and stars and the radius of the disc; these quantities are predicted by GALFORM. The BR06 SF law contains no free parameters once it has been calibrated against observations. We adopt a value for the normalization of the surface density of SF, $\nu_{\rm SF}$ (see Table 1), that is two-thirds the value used by Lagos et al., but which is still within the observational uncertainty. By adopting a lower value for $\nu_{\rm SF}$ we achieve a much steeper evolution of the quiescent star formation rate (SFR) density in the $z \simeq 0$–2 universe (see the next paragraph). We do this to improve the match to the observed quasar LF, since the AGN activity is directly linked to the SF in the host galaxy.

As shown by Lagos et al., the adoption of the BR06 SF law changes the SF history predicted by the model. This is due to a much weaker dependence of the effective SF time-scale on redshift with the new SF law, compared with that displayed by the Bower et al. model. This leads to the buildup of larger gas reservoirs in discs at high redshift, resulting in more SF in bursts. This is shown in the top panel of Fig. 1, which compares the predictions of the F11b (solid lines) and F11a (dashed lines) models for the total SFR density and the SFR density in the quiescent and burst SF modes. Fig. 1 shows that the individual SF modes change substantially on using the BR06 SF law. The quiescent SF mode in the F11b model is suppressed by almost an order of magnitude above $z \sim 3$ compared to the F11a model. This has an impact on the buildup of BH mass, by changing the amount of mass brought in through

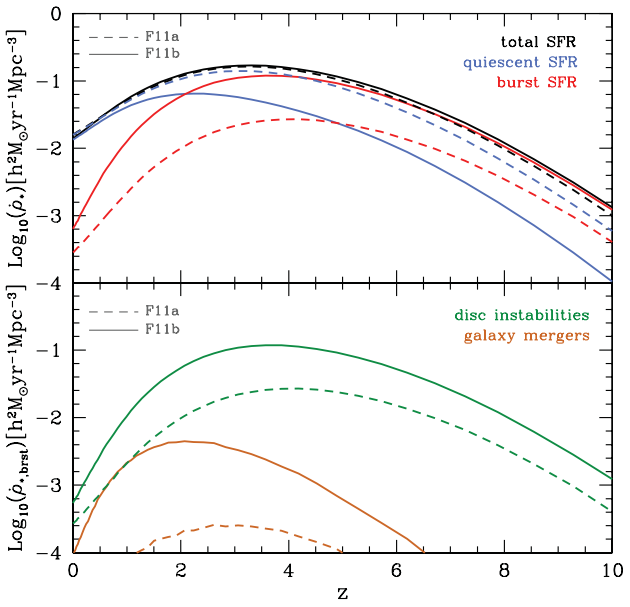

Figure 1. Top panel: the cosmic history of the total SFR density (black lines) in the F11b (solid line) and F11a (dashed line) models. Also shown is the contribution from each SF mode, namely burst and quiescent SF (red and blue, respectively), to the total SFR density. Bottom panel: the contribution of disc instabilities (green lines) and galaxy mergers (orange lines) to the cosmic history of SFR density in the burst SF mode for the F11b (solid lines) and F11a (dashed lines) models.

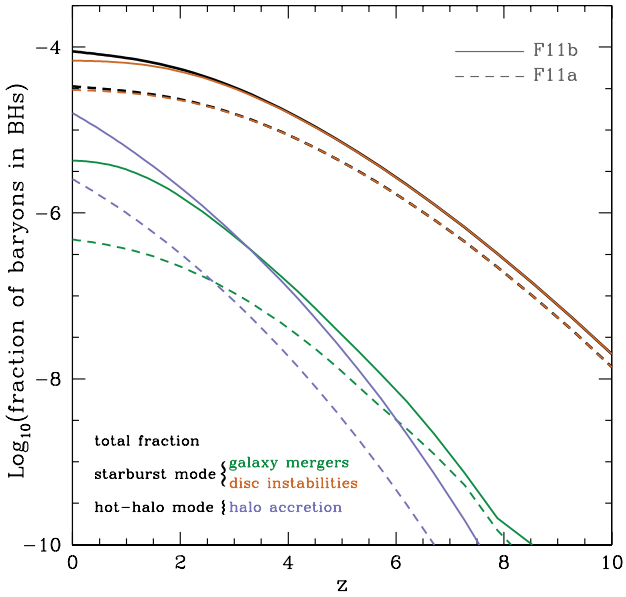

Figure 2. The fraction of baryons locked in BHs as a function of redshift (black lines) for the F11b (solid lines) and F11a (dashed lines) models. The plot also shows the contribution of each accretion channel, namely disc instabilities (orange), galaxy mergers (green) and quasi-hydrostatic halo accretion (blue), to the total fraction of baryons in BHs.

the starburst mode. The enhancement of the burst mode is further demonstrated in the bottom panel of Fig. 1, where we show the SFR density history in bursts distinguishing between those triggered by disc instabilities and galaxy mergers for the F11a and F11b models. Both channels show a significant increase in SFR density in the F11b model. Since the BHs in our model grow in bursts of SF, we expect this enhancement to have a significant impact on the evolution of BH mass.

In Fig. 2, we plot the fraction of baryons locked up in BHs as a function of redshift. We distinguish between the different accretion channels: disc instabilities, galaxy mergers and accretion from hot gas haloes in quasi-hydrostatic equilibrium. Fig. 2 demonstrates that both the disc instability and galaxy merger channels are enhanced when using the BR06 SF law, resulting in differing evolution of the total BH mass. The F11b model also predicts different hot-halo mode accretion. This is due to the different values of the $\epsilon_{\rm kin}$ and $f_{\rm Edd}$ parameters used in F11a (note that halo accretion is completely independent of SF in galactic discs). We note that, as shown in Fig. 2, the growth of BHs is dominated by accretion of cold gas during disc instabilities. Hence, secular processes in galaxies are responsible for building most of the BH mass, while galaxy mergers become an important channel only when they occur between galaxies of similar mass (i.e. major mergers). This merely reflects the importance of the disc-instability channel in triggering starbursts and building the stars in galactic bulges in our model.

The change in the total fraction of baryons locked up in BHs seen in the F11b model affects mostly the most massive BHs. This is illustrated in Fig. 3, where we plot the BH mass function (BHMF) predicted by the F11a and F11b models at $z = 0$. In addition, we show two more variants of the Bower et al. model to further demonstrate the dependance of BH mass on the parameters that influence the evolution of BHs ($f_{\rm BH}$, $f_{\rm Edd}$ and $\epsilon_{\rm kin}$; these variants consider the BR06 SF law). Fig. 3 shows that the parameters related to the AGN feedback mechanism, $f_{\rm Edd}$ and $\epsilon_{\rm kin}$, shape the high-mass end of the BHMF. This is because the feedback mechanism is most

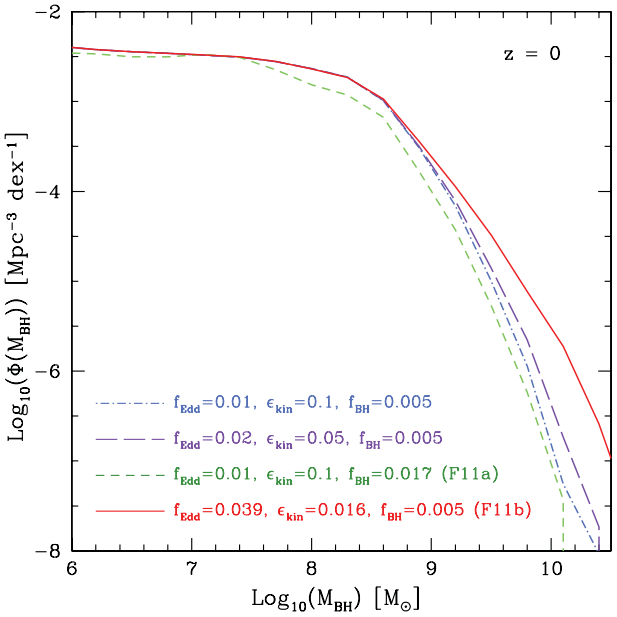

Figure 3. The mass function of BHs at $z = 0$ for four different variations of the Bower et al. model, including the fiducial model F11b, indicated by the legends.

efficient in the most massive haloes and therefore it affects the most massive BHs. When considering the first set of paramaters ($f_{BH} = 0.005, f_{Edd} = 0.01$ and $\epsilon_{kin} = 0.1$), we find the following interesting result. Modest values of the f_{Edd} parameter mean that less accretion energy is available for heating the gas in the halo. This should in principle allow BHs to grow to significant masses. However, the model suggests that by choosing a relatively high (e.g. 0.1) average jet kinetic efficiency, ϵ_{kin}, during the hot-halo mode a very strong feedback is achieved and, therefore, the BH growth is significantly suppressed. Hence, to enhance the growth of massive BHs, we need to decrease the jet efficiency by a significant factor and allow more accretion energy to heat the gas in the host halo (see e.g. the second variant of the Bower et al. model with $f_{Edd} = 0.02$ and $\epsilon_{kin} = 0.05$, and the fiducial model F11). Furthermore, by changing the value of the f_{BH} parameter, we can extend the dynamical range of BH masses built in the model since a high value of f_{BH} implies more gas is being deposited into the BHs. This produces an overall shift of mass function along the BH mass axis. The properties of the BHMF will be explored more thoroughly in the next section.

Finally, we note that the total SFR density history in Fig. 1 remains essentially unchanged when using the BR06 SF law, as reported by Lagos et al. This implies that properties such as stellar masses and galaxy luminosities are fairly insensitive to the choice of SF law. Indeed, Lagos et al. show that the predictions for the $z = 0$ b_J- and K-band LFs on using the BR06 SF law are very similar to those of the original Bower et al. model and reproduce the observations very closely. Since we have changed the normalization of the SF law, we recheck that our model still reproduces the observed LF of local galaxies, as shown in Fig. 4. Hence, we are confident that we are building an AGN model within a realistic galaxy formation model. In the following sections, we will describe the AGN model extensively and present its predictions for the evolution of BH mass, spin and AGN abundances. When running GALFORM, the parameters relating to the galaxy formation part of the code (including those listed in Table 1) are kept fixed. The only free parameters in our model are strictly related to the accreting BHs and do not affect the evolution of their host galaxies. These parameters, which are tuned to achieve a good fit to the available observational data on AGN LFs, will be described throughout the text and finally summarized in Table 3 (see Section 4.1).

3 THE EVOLUTION OF BH MASS

BH demographics have been the topic of many studies in the past decade mainly because of the tight correlations between the properties of BHs and their host stellar spheroids. These correlations take various forms, relating, for example, the mass of the BH to the mass of the galactic bulge (the M_{BH}–M_{Bulge} relation: Magorrian et al. 1998; McLure & Dunlop 2002; Marconi & Hunt 2003; Häring & Rix 2004), or to the stellar velocity dispersion (the M_{BH}–σ_* relation: Ferrarese & Merritt 2000; Gebhardt et al. 2000; Tremaine et al. 2002). These remarkable and unexpected correlations suggest a natural link between the BH evolution and the formation history

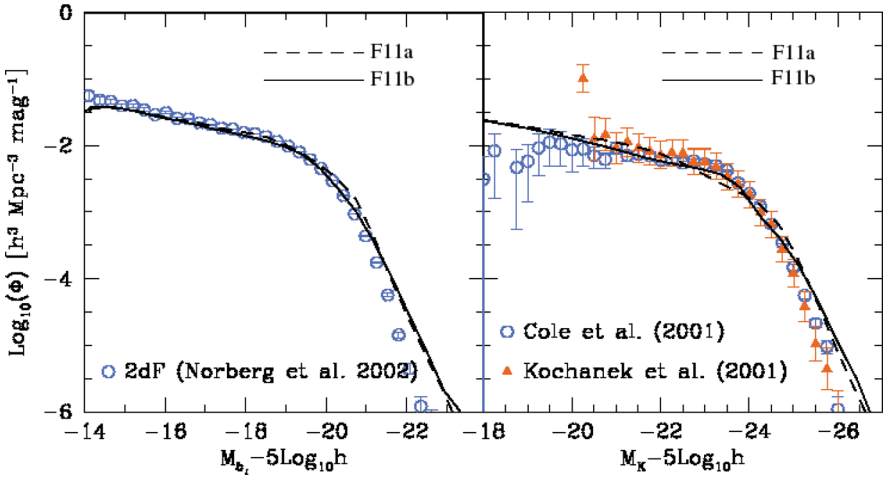

Figure 4. The LF of galaxies in the local Universe. The left-hand panel compares the predictions of the F11b (solid line) and F11a (dashed line) models for the b_J-band LF with the observational determination from the 2dF galaxy redshift survey by Norberg et al. (2002). Similarly, the right-hand panel shows the predictions of the models (line style as before) for the K-band LF compared to the observational determinations by Cole et al. (2001) and Kochanek et al. (2001). The theoretical predictions from both models include dust extinction.

© 2011 The Authors, MNRAS **419,** 2797–2820
Monthly Notices of the Royal Astronomical Society © 2011 RAS

Table 2. Summary of the local BH mass densities found in this and previous studies (assuming $h = 0.7$). Densities are shown for BHs in disc (S, S0) and elliptical (E) galaxies, and for the global BH population in our model (tot).

Study	$\rho_{\rm BH}$(S, S0)^a	$\rho_{\rm BH}$(E)^b	$\rho_{\rm BH}$(tot)^c
This study	0.72	4.70	5.42
Graham et al. (2007)	$0.95^{+0.49}_{-0.49}$	$3.46^{+1.16}_{-1.16}$	$4.41^{+1.67}_{-1.67}$
Shankar et al. (2004)	$1.1^{+0.5}_{-0.5}$	$3.1^{+0.9}_{-0.8}$	$4.2^{+1.1}_{-1.1}$
Marconi et al. (2004)	1.3	3.3	$4.6^{+1.9}_{-1.4}$
Fukugita & Peebles (2004)	$1.7^{+1.7}_{-0.8}$	$3.4^{+3.4}_{-1.7}$	$5.1^{+3.8}_{-1.9}$

^{a,b,c}In units of $10^5\,{\rm M}_\odot\,{\rm Mpc}^{-3}$.

of galaxies. The manifestation of this link could be associated with AGN activity triggered during the buildup of BHs.

The observed $M_{\rm BH}$–$M_{\rm Bulge}$ relation can be used to estimate the mass of a BH. Several authors have utilized this technique to estimate the BHMF in the local Universe (Yu & Tremaine 2002; Marconi et al. 2004; Shankar et al. 2004), using scaling relations such as $M_{\rm BH}$–σ_* and $M_{\rm BH}$–$L_{\rm Bulge}$. Based on these MFs, and others inferred by independent studies, the total BH mass density in the local Universe has been estimated to be in the range $\rho_{\rm BH} = (4.2\text{–}5.1) \times 10^5\,{\rm M}_\odot\,{\rm Mpc}^{-3}$ (see Table 2 for a list of local BH mass density estimates).

The global BHMF predicted by the model for all BHs at $z = 0$ is shown in Fig. 5(a). The predicted MF is almost constant over the mass range 10^6–$10^8\,{\rm M}_\odot$. For higher masses, the MF decreases steeply with increasing mass. Our predictions are compared to the local MFs estimated by Marconi et al. (2004) and Shankar et al. (2004) using the BH mass scaling relations in local galaxies. The predicted and observed MFs are in good agreement for low-mass BHs and disagree moderately for the 10^9–$10^{10}\,{\rm M}_\odot$ BHs. The overall BH mass density at $z = 0$ predicted by our model amounts to $\rho_{\rm BH} = 5.42 \times 10^5\,{\rm M}_\odot\,{\rm Mpc}^{-3}$ and is consistent with that estimated by Marconi et al. and other authors (Shankar et al. 2004; Graham et al. 2007; Yu & Lu 2008; see Table 2). Note that, when calculating the BH mass density, we do not take into account mass-loss due to gravitational wave emission from BH mergers (Menou & Haiman 2004).

In Fig. 5(a), we also show the contribution to the MF from BHs hosted by elliptical galaxies. We classify as ellipticals all the galaxies whose bulge luminosity contributes more than 60 per cent to their total $b_{\rm J}$-band luminosity (Cole et al. 2000; Parry, Eke & Frenk 2009). The rest of the galaxies are classified as spiral or S0 galaxies. As expected, the BHs in the elliptical galaxies dominate the high-mass end of the MF. Hence, the most massive BHs inhabit the centres of the most massive systems in our simulations. By contrast, spiral and S0 galaxies contribute only to the low-mass end of the MF.

The growth of BHs in spiral and elliptical galaxies is driven in principle by different channels. In spiral galaxies, BHs grow mainly via accretion during the starburst mode, whereas the massive BHs in ellipticals grow via accretion during the hot-halo mode or BH–BH mergers. Evidently, the evolution of the BHMF is shaped by these growth channels. This is illustrated in Fig. 5(b), where we show the global BHMF for eight redshift snapshots of the Millennium simulation in the range $z = 0$–6.2. The buildup of the BHs populating the low-mass end takes place at high redshifts and is complete at $z \sim 2$. Since $z \sim 2$, the amplitude of the BHMF remains almost unchanged, indicating that 10^6–$10^8\,{\rm M}_\odot$ BHs in spiral galaxies today were already in place by $z \sim 2$, which follows from the spheroid being older than the disc. The growth of these lower mass BHs is dominated by accretion during the starburst mode. Interestingly, other models that match the LF of accreting BHs indicate a different evolution for the BHMF (see e.g. Hopkins et al. 2007; Shankar et al. 2009). These studies suggest that the most massive BHs ($M_{\rm BH} > 10^9\,{\rm M}_\odot$) are in

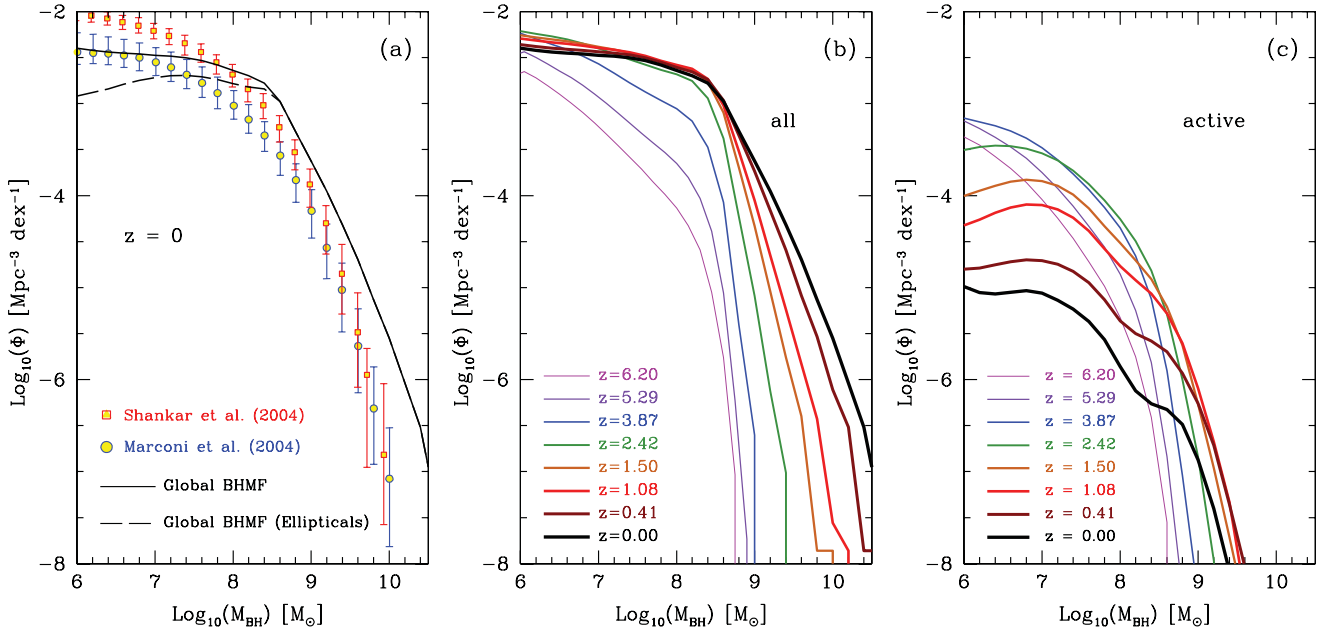

Figure 5. Panel (a): the global BHMF at $z = 0$ predicted by our model (black solid line). The observationally estimated MFs are taken from Marconi et al. (2004, blue filled circles) and Shankar et al. (2004, red filled squares). The error bars correspond to $\pm 1\sigma$ uncertainties. Panel (b): the evolution of the global BHMF with redshift. Panel (c): the MF of actively growing BHs (BHs accreting in the thin-disc regime: $\dot{m} > 0.01$) calculated at different redshifts as indicated in the legend.

© 2011 The Authors, MNRAS **419,** 2797–2820
Monthly Notices of the Royal Astronomical Society © 2011 RAS

place already at $z \sim 3$. In contrast, the abundance of less massive BHs is found to decline quickly as redshift increases. This evolutionary trend is an indication of the downsizing of BH mass, a topic that will be discussed extensively in Section 5.5. For the moment, we note that the fact that our model matches the observational constraints for the BH mass (and also the LFs presented in Section 5) suggests that the models of the BHMF that indicate downsizing do not provide a unique interpretation of the observational data.

The fact that the space density of this population does not evolve since $z \sim 2$ suggests that the starburst mode becomes less significant in the low-redshift universe. None the less, the buildup of BH mass continues below $z \sim 2$, where we witness the buildup of the most massive BHs in the universe. These BHs grow mainly during the hot-halo accretion mode or via mergers with other BHs (see Fanidakis et al. 2011). The different physics governing these channels results in a different slope for the high-mass end of the BHMF giving rise to a strong break at $\sim 5 \times 10^8\,\mathrm{M}_\odot$. Yet, the buildup of these BHs is not very efficient. This becomes clearer when we consider the evolution of the BH mass density. For instance, in the redshift range $z = 6.2$–2, the BH mass density increases by a factor of $\sim$25, as a result of the intense accretion during the starburst mode in the high-redshift universe. By contrast, from $z = 2$ until $z = 0$ the density only doubles, mainly because the 10^9-$\mathrm{M}_\odot$ BHs become more numerous.

To gain more insight into the rapidly evolving population of accreting BHs, we consider the evolution of actively growing BHs, namely those BHs that accrete at relatively high rates (greater than 1 per cent of their Eddington accretion rate, see Section 4). In Fig. 5(c), we show the MF of actively growing BHs in our model and its evolution with redshift. The most striking characteristic of the MF is the dramatic change in shape and normalization with redshift. At $z = 0$, the MF has a maximum at $\sim 10^7\,\mathrm{M}_\odot$ below which the accreting BHs have an almost constant space density (see also results from the SDSS, Heckman et al. 2004). A significant bend in the curve at $\sim 10^8\,\mathrm{M}_\odot$ is seen, a feature that remains evident up to $z \sim 2$–3. At higher redshifts, the space density of BHs drops substantially, establishing that accreting BHs with masses $\gtrsim 10^9\,\mathrm{M}_\odot$ are very rare objects. These BHs are accreting at relatively low rates and therefore power-faint AGN activity (see discussion in Section 4.4).

As illustrated in Fig. 5(c), the low-mass end of the MF evolves differently with redshift compared to the high-mass end. Its amplitude increases significantly with increasing redshift, suggesting that the space density of low-mass accreting BHs was higher at earlier epochs. This evolution is monotonic up to $z \sim 2.5$ above which the amplitude of the low-mass end starts to decline modestly. The high-mass end shows a similar increase towards higher amplitudes with increasing redshift; however, the decline appears earlier, almost at $z \sim 1$. This is because the high-mass end forms much later than the low-mass end in our model.

Greene & Ho (2007) calculated the BHMF of type-1 active galaxies with $z < 0.35$ in the SDSS Data Release 4 sample (DR4, Adelman-McCarthy et al. 2006) and found that it displays a strong break at $\sim 10^{6.6}\,\mathrm{M}_\odot$, and declines steeply towards both higher and lower mass. Their results suggest that the mass of typical active BHs in the local Universe is $\simeq 10^7\,\mathrm{M}_\odot$ and has a characteristic accretion rate of $\sim$10 per cent of the Eddington value. More recently, Schulze & Wisotzki (2010) presented the active BHMF of local ($z < 0.3$) galaxies from the ESO/Hamburg survey (Wisotzki et al. 2000) and found good agreement with the BHMFs in AGNs from the SDSS. Their results indicate that the fraction of active galaxies decreases rapidly for high BH masses. By comparing with BHMFs of higher redshifts (Vestergaard et al. 2008; Vestergaard & Osmer 2009) Schulze & Wisotzki argue that less massive BHs are currently active, a conclusion that suggests an anti-hierarchical growth of BHs in the Universe.

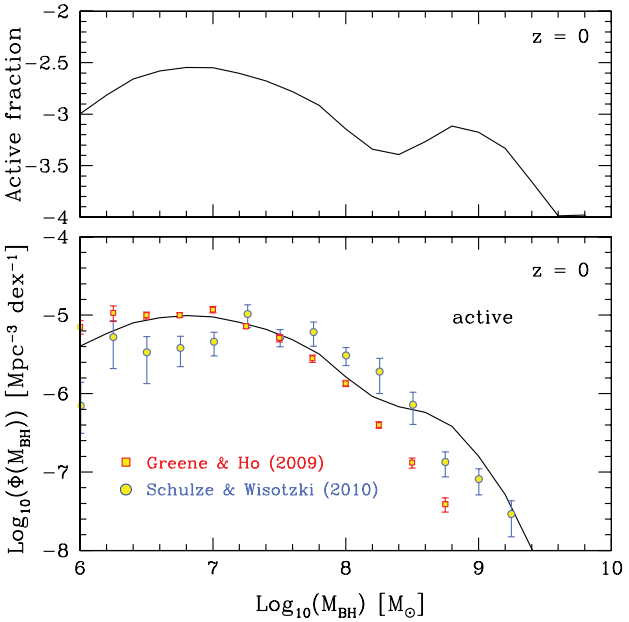

Figure 6. Bottom panel: comparison of recent observational estimations of the active BHMF ($\log_{10}\lambda_{\mathrm{Edd}} > -2$) at $z = 0$ with the predictions of our model. Data are taken from Greene & Ho (2009, revised version) and Schulze & Wisotzki (2010). Top panel: the fraction of active galaxies at $z = 0$ as a function of BH mass.

To test these hypotheses, we compare our predictions for the BHMF of type-1 active galaxies in the local Universe with the observational estimate from Schulze & Wisotzki (2010) and Greene & Ho (2007). The comparison is shown in Fig. 6 where we plot the $z = 0$ active BHMF and those from the SDSS DR4 and ESO/Hamburg surveys (bottom panel). We note that Greene & Ho recently pointed out that a factor of $1 + z$ was omitted in their calculation of the maximum volume weights. Here we present their corrected data as published in Greene & Ho (2009). To calculate the active BHMF, we consider all the BHs that accrete above 1 per cent of the Eddington limit (see Section 4). We also show the fraction of active BHs in the same redshift range, calculated as $f_{\mathrm{act}}(M_{\mathrm{BH}}) = \Phi(M_{\mathrm{BH}})_{\mathrm{act}}/\Phi(M_{\mathrm{BH}})$ (top panel). Our predictions are in good agreement with both observational data sets. The model reproduces the turnover at $\sim 10^7\,\mathrm{M}_\odot$ suggested by the observations and the characteristic decline at higher BH masses, which is associated with the decline of accreting gas in massive galaxies. The decrease in the BHMF at lower masses in our model is due to the luminosity limit adopted to reproduce the selection applied to the observational samples.

The fraction of active galaxies in the top panel of Fig. 6 indicates that an appreciable fraction ($\sim$0.3 per cent) of BHs with masses of $\sim 10^7\,\mathrm{M}_\odot$ are currently active. A significantly lower active fraction is found for higher BH mass, as suggested by the steep slope above $\sim 10^7\,\mathrm{M}_\odot$. Our model further predicts a characteristic 'bump' at $\sim 10^9\,\mathrm{M}_\odot$. This bump is related to the tail of BHs accreting during the hot-halo mode at rates higher than 1 per cent of the Eddington value.

We finally note that the fraction of active BHs determines the BH 'duty cycle' (Shankar 2009; Shankar, Weinberg & Miralda-Escudé 2009). The BH duty cycle in our model strongly depends on the

accretion time-scale, which is proportional to the dynamical time-scale of the host bulge (see Section 4.1). The proportionality factor, f_q, determines the normalization of the duty cycle and therefore the fraction of active BHs in a given mass bin. As we shall see in the following section, we set the value of f_q to 10 by fitting the LFs of AGNs. The value obtained from the best fit naturally reproduces also the duty cycle of BHs as shown in Fig. 6, and therefore the model can provide a consistent framework within which active BHs can be studied.

The active BHMF in Fig. 6 and its evolution shown in Fig. 5(c) indicates an evolutionary scenario for the actively growing BH population in our model which mirrors the hierarchical growth of their host galaxies. At high redshifts, the accretion activity is dominated by 10^6-$M_\odot$ BHs. Accretion on to these BHs results in the fast buildup of 10^7–10^8 $M_\odot$ BHs at $z \sim 4$–6, as implied by the increase in the space density of these BHs in Fig. 5(c). From $z \sim 4$, the accretion activity gradually shifts to the 10^7–10^8 $M_\odot$ BH population. This is mainly because more higher mass galactic systems are now in place and thus disc instabilities and galaxy mergers contribute to the growth of higher mass BHs compared to earlier epochs. Accretion on to these BHs leads to the fast buildup of the $>10^8$ $M_\odot$ BH population, which will later be promoted via accretion during the hot-halo mode and BH–BH mergers to the most massive BHs in our model.

Eventually, the different physical processes that drive the evolution of BH mass result in a tight correlation between the BH and host galaxy mass, as shown in Fig. 7(a). In this figure, we plot the median of the M_{BH}–M_{Bulge} distribution (solid lines) and the associated 25–75 and 5–95 percentiles (shaded areas) at $z = 0$, 1.08, 2.42 and 6.2. Surprisingly, a well-defined M_{BH}–M_{Bulge} relation is established already at $z = 6.2$. This is a consequence of the tight correlation between the accretion of gas and the SF that shapes the mass of the host bulges. Even when the growth of BHs is dominated by other processes, which are only indirectly linked to the SF, the tight correlation remains. For example, the mass of the most massive BHs correlates with the mass of their host bulge because feedback energy regulates the cooling flows that supply gas for SF and therefore the growth of the bulge mass.

In a hierarchical universe, the most massive galactic bulges are usually found in the most massive DM haloes. Since the galaxies with the most massive bulges (namely the passive ellipticals) host the most massive BHs and those with less massive bulges (namely the spiral and S0 galaxies) host the low-mass BHs, it is natural to assume that there must be a similar hierarchy in the halo environments where these BHs can be found. Indeed, when plotting the relation between the BH mass and the mass of the host halo, we find a tight correlation. This is illustrated in Fig. 7(b) where we show the median of the M_{BH}–M_{Halo} distribution (solid lines) and its 10–90 percentiles (shaded areas) for different redshifts. Remarkably, the physical processes that shape the M_{BH}–M_{Bulge} relation in our model give rise to a well-defined link between these two quantities at all redshifts. The M_{BH}–M_{Halo} relation can be parametrized by a power law with a very steep slope that becomes shallower above $M_{Halo} \simeq 10^{12}$–10^{13} $M_\odot$.

The different slopes indicate the different efficiency with which BHs grow in haloes of different mass.

Low-mass haloes are very efficient in growing BHs since in these environments the gas cooling is not suppressed. As a consequence, BHs double their mass remarkably quickly resulting in a steep slope for the M_{BH}–M_{Halo} relation. The fast BH mass buildup slows down when hot gas in the host haloes enters the quasi-hydrostatic regime. In this case, AGN feedback suppresses the cooling flows that provide fresh cold gas to the galaxies and thus accretion through the starburst mode is reduced. In the quasi-hydrostatic regime, gas accretion during the hot-halo mode and mergers dominate the BH mass buildup. The hot-halo mode is, however, characterized by low accretion rates (see Section 4.3). BH mergers rather than adding new baryons to the BHs only redistribute the BH mass. Therefore, the BH mass

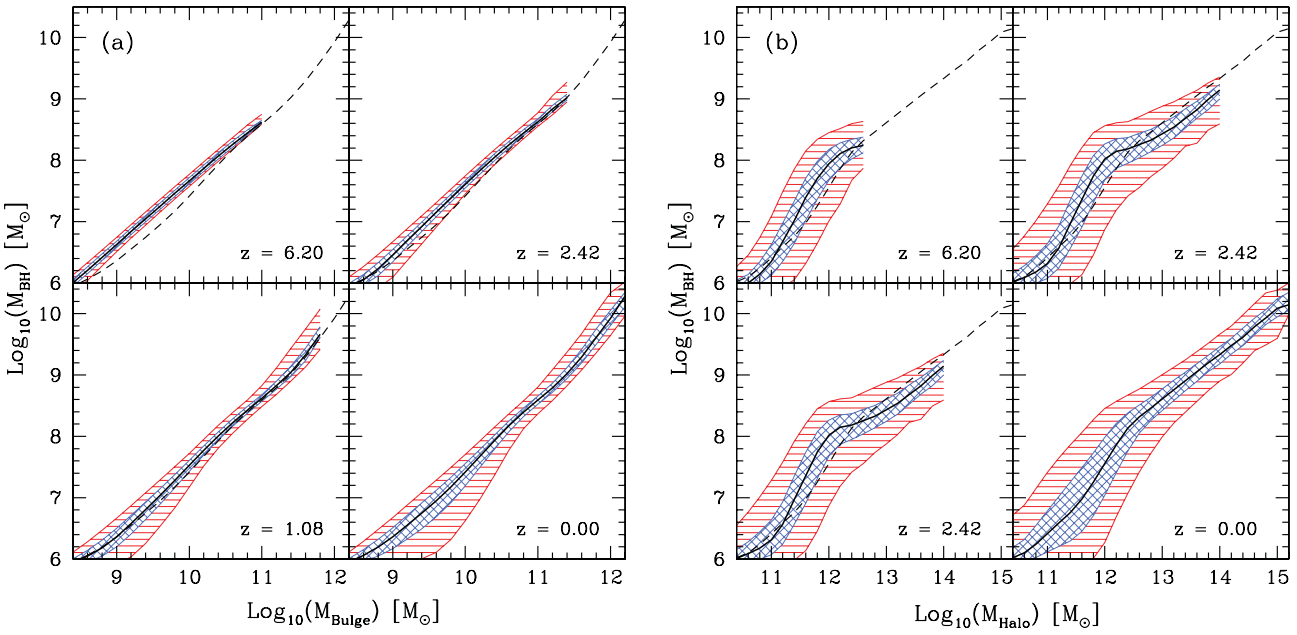

Figure 7. The median of the (a) M_{BH}–M_{Bulge} and (b) M_{BH}–M_{Halo} distributions predicted by our model (solid lines). Predictions are shown for $z = 0$, 1.08, 2.42 and 6.20 as indicated in the legend. The shaded areas indicate the 5–95 (red) and 25–75 (blue) percentile ranges of the distributions. The median of the (a) M_{BH}–M_{Bulge} and (b) M_{BH}–M_{Halo} distributions at $z = 0$ is shown in every panel for reference (dashed black line).

buildup slows down, establishing haloes of $M_{\rm Halo} \sim 10^{13}$–$10^{15}\,{\rm M}_{\odot}$ as environments where BH growth is not very efficient.

Given the efficiencies that characterize the two different regimes in Fig. 7(b), we expect to find the brightest AGNs in the 10^{11}–$10^{13}\,{\rm M}_{\odot}$ haloes. This is in contrast with the common expectation that the brightest quasars should be found in the most massive haloes.

4 BLACK HOLE SPINS, ACCRETION EFFICIENCIES AND DISC LUMINOSITIES

The ability of an accretion disc to produce electromagnetic radiation is attributed to gravity, yet its radiative efficiency is controlled primarily by the properties of the gas. When the gas settles itself into a thin, cool, optically-thick accretion disc (Shakura & Sunyaev 1973), the efficiency of converting matter into radiation can reach 32 per cent. When the flow is characterized by very high, Eddington accretion rates, such extremely luminous discs can power 10^{47}–$10^{48}\,{\rm erg\,s^{-1}}$ quasars. In the next sections, we describe how we model the physics of accretion flows and present predictions for the most fundamental properties characterizing the accreting BH systems in our model.

4.1 Calculation of the disc luminosity

The first important property that we can calculate in our model is the physical accretion rate, $\dot{M}$, on to a BH. This is defined as

$$\dot{M} = \frac{M_{\rm acc}}{t_{\rm acc}}, \tag{1}$$

where $M_{\rm acc}$ is the total accreted mass and $t_{\rm acc}$ is the accretion time-scale. The accretion time-scale is assumed to be directly linked to the dynamical time-scale of the host bulge, $t_{\rm Bulge}$, through the relation

$$t_{\rm acc} = f_{\rm q} t_{\rm Bulge}. \tag{2}$$

$f_{\rm q}$ is a free parameter and its value is fixed to 10 by fitting the predictions of our model for the quasar LF to the observations (see Section 5.2). This is a rather simple assumption. In reality, it is expected that the accretion time-scale is likely to be complex and almost certainly not constant throughout the process of accretion. We choose this dependence for simplicity in the absence of a robust understanding of the accretion process in AGNs.

A second important quantity associated with every BH is the Eddington luminosity,

$$L_{\rm Edd} = \frac{4\pi G M_{\rm BH} m_{\rm p} c}{\kappa} = 1.4 \times 10^{46} \left(\frac{M_{\rm BH}}{10^8\,{\rm M}_{\odot}}\right) {\rm erg\,s^{-1}}, \tag{3}$$

where $\kappa \sim 0.3\,{\rm cm^2\,g^{-1}}$ is the electron scattering opacity and $m_{\rm p}$ is the proton mass. The Eddington luminosity has an associated accretion rate which is expressed as

$$\dot{M}_{\rm Edd} = L_{\rm Edd}/\epsilon c^2, \tag{4}$$

where ϵ is the accretion efficiency and c is the speed of light. This is the accretion rate at which the BH radiates at the Eddington luminosity. The accretion efficiency is assumed to be determined by the spin of the BH (Novikov & Thorne 1973) and its value is calculated as in Fanidakis et al. (2011). It is convenient in our analysis to express the physical accretion rate in units of the Eddington accretion rate, $\dot{m} \equiv \dot{M}/\dot{M}_{\rm Edd}$, in order to introduce a dependence on the BH mass.

The accretion rate has a dramatic impact on the geometry and radiative properties of the accretion disc (see Done, Gierliński & Kubota 2007). In our model, we consider two distinct accretion modes separated at a rate of 1 per cent of $\dot{M}_{\rm Edd}$. In the first state, for $\dot{m} \geq 0.01$, we assume that the gas forms an accretion disc whose physics is adequately described by the radiatively efficient thin-disc model of Shakura & Sunyaev (1973). The bolometric luminosity of a thin disc, $L_{\rm bol}$, is linked to $\dot{M}$ through the standard expression

$$L_{\rm bol} = \epsilon \dot{M} c^2. \tag{5}$$

When the accretion becomes substantially super-Eddington ($L_{\rm bol} \geq \eta L_{\rm Edd}$), the bolometric luminosity is limited to (Shakura & Sunyaev 1973)

$$L_{\rm bol}(\geq \eta L_{\rm Edd}) = \eta[1 + \ln(\dot{m}/\eta)] L_{\rm Edd}, \tag{6}$$

where η is an ad hoc parameter, which we choose equal to 4, that allows a better modelling of the bright end of the LF (see Section 5). However, we do not restrict the accretion rate if the flow becomes super-Eddington.

The second accretion state, with $\dot{m} < 0.01$, is associated with sub-Eddington accretion flows with very low density. For such low accretion rates, the gas flow is unable to cool efficiently since radiative cooling does not balance the energy generated by viscosity. Thus, the viscous energy is trapped in the gas and ultimately advected into the hole. This type of accretion is known as an ADAF (Rees 1982; Narayan & Yi 1994; Abramowicz et al. 1995). ADAFs have a number of distinct properties, some of which will be essential for the analysis in later sections (see Sections 5 and 6). For example, for an ADAF around a BH, only a fraction of the standard accretion luminosity, $L = \epsilon \dot{M} c^2$, is emitted as radiation. The remainder of the viscously dissipated energy is stored in the gas as entropy, resulting in hot flows with almost virial temperatures. We note that, as shown by Ichimaru (1977), the ions and the electrons in an ADAF are not thermally coupled and, thus, reach different temperatures. This two-temperature virialized plasma flow is optically thin and, for high values of the Shakura–Sunyaev viscosity parameters α ($\sim$0.1–0.3, Narayan & Yi 1994), it acquires a quasi-spherical geometry around the BH, which resembles spherical Bondi accretion. However, the accretion is entirely due to dissipation via viscous forces rather than gravity. The bolometric luminosity of the flow in this case is equal to the luminosity emitted by the various cooling processes. For $\dot{m} \lesssim 10^{-3}\alpha^2$, it is only due to Comptonization, whereas for $\dot{m} \gtrsim 10^{-3}\alpha^2$, the cooling is split between Compton and synchrotron emission.

As the accretion rate increases, the emission of the energy produced by the viscous processes in the gas becomes more efficient. Above some critical accretion rate, $\dot{m}_{\rm c}$, the radiative efficiency of the gas is so high that the flow cools down to a thin disc. The critical $\dot{m}$ is independent of BH mass but depends strongly on the viscosity, $\dot{m}_{\rm c} \simeq 1.3\alpha^2$. Taking $\dot{m}_{\rm c} = 0.01$ implicitly fixes the value of α to be 0.087 for all the ADAFs in our model. The luminosity of the flow for the two different regimes is then given by Mahadevan (1997),[2]

$$L_{\rm bol,ADAF} = \begin{cases} (0.44\dot{m}/0.01)\epsilon \dot{M} c^2, & \dot{m} > 7.5 \times 10^{-6} \\ 6.3 \times 10^{-5} \epsilon \dot{M} c^2, & \dot{m} \lesssim 7.5 \times 10^{-6}. \end{cases} \tag{7}$$

The expression in the brackets in the first regime shows how much less efficient the cooling is in an ADAF compared to the standard efficiency ϵ of a thin disc. For example, at $\dot{m} = 0.01$, an ADAF is

[2] In these expressions, we have replaced the β parameter from the original Mahadevan scaling relations with the relation $\alpha \approx 0.55(1 - \beta)$ (Hawley, Gammie & Balbus 1995).

Table 3. Summary of the values of the free parameters in the AGN model.

Study	f_q [a]	α [b]	η	$\dot{m}_c$ [c]
F11b (this study)	10	0.087	4	0.01
F11a (Fanidakis et al. 2011)	4	0.1	1	0.01

[a]Proportionality factor in the accretion time-scale in equation (2).
[b]Shakura–Sunyaev viscosity parameter of the accretion disc.
[c]Critical accretion rate above which transition from an ADAF to a thin disc occurs.

characterized by an accretion efficiency of 0.44ϵ, less than approximately half as luminous as a thin disc.

Finally, in Table 3, we summarize the most important parameters in the AGN model. The values of $\dot{m}_c$, η and α are constrained according to the discussion in the previous paragraphs. In practice, only f_q has a significant impact on the model predictions and can be adjusted to match the observed AGN LF. The parameter η affects only the slope of the exponential cut-off of the LF and hence influences a small fraction of the AGN population (see Fig. 17 shown later).

4.2 BH spins and accretion efficiencies

In Fanidakis et al. (2011), we describe in detail how we model the evolution of BH spin. In brief, the spins of BHs change during gas accretion (whenever a starburst is triggered by disc instabilities or galaxy mergers in the starburst mode or during the cooling of gas from the hydrostatic halo in the hot-halo mode) and mergers with other BHs. Fig. 8 shows the evolution with redshift of the BH spin distribution for two different cases. First, we show the median spin weighted by BH mass (upper panel) for all the BHs with $M_{BH} > 10^6\,M_\odot$. The median shows an approximately constant trend from $z = 6$ to 2 and has a well-defined value of ∼0.25. In this redshift range, BHs grow predominantly during disc instabilities (see Fig. 2) and therefore the evolution of BH spin is governed by accretion. As shown by Fanidakis et al., accretion of gas results in low spins with a typical value of $a \sim 0.2$ (under the assumption that the gas is fed chaotically on to the BH). Hence, as indicated also by the different percentiles in the plot (shaded regions), the bulk of BHs acquire low spins.

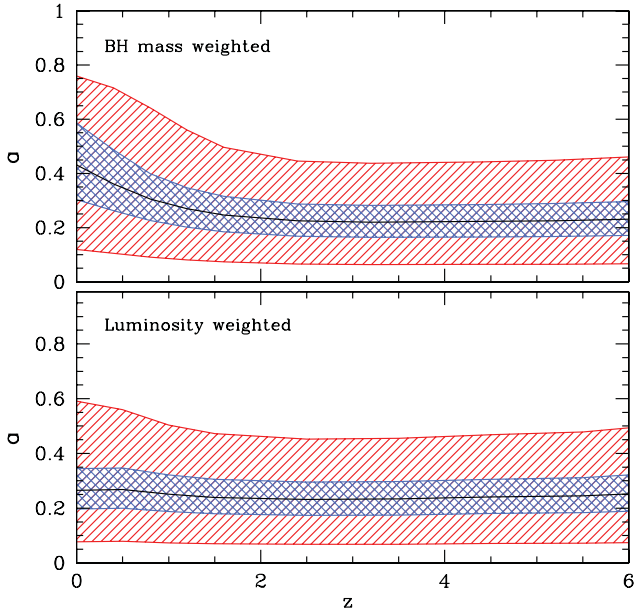

Figure 8. Top panel: the median BH spin weighted by BH mass as a function of redshift for BHs in our model (solid black line). Bottom panel: the median BH spin weighted by disc luminosity for all AGNs accreting in the thin-disc regime (solid black line). Also shown in each panel are the 10–90 (red shaded areas) and 25–65 (blue shaded areas) percentile ranges of the BH spin distribution.

Below $z = 2$, BH mergers start to become an important channel for growing the BH mass, especially when they are between BHs of similar mass, as expected following a major galaxy merger. However, this is a characteristic only of the most massive BHs in our model ($M_{BH} > 10^9\,M_\odot$). Mergers between those BHs tend to increase the spin of the final remnant to values $a > 0.7$ (see also Baker et al. 2007; Berti & Volonteri 2008; Fanidakis et al. 2011). Eventually, at low redshifts, BH mergers give rise to a population of rapidly rotating BHs. The appearance of these BHs, as indicated also by the different percentiles below $z = 2$, increases the range of the predicted spins up to values of 0.998 (the range of spin values in the distribution of BHs in Fig. 8 is smaller because we show only up to the 90th percentile of the data).

The environmental dependence of BH spins is illustrated in Fig. 9, where we see that the most rapidly rotating BHs populate the centres of the most massive DM haloes. In contrast, slowly rotating BHs are found preferentially in low-mass haloes (with some scatter). This is a consequence of the correlation between BH mass and spin in our model: rapidly rotating BHs have masses $\gtrsim 10^9\,M_\odot$. These BHs are hosted by massive elliptical galaxies in our model. In a forthcoming paper, we will explore the clustering properties of rapidly rotating BHs and show quantitatively the environmental dependence of BH spin.

In contrast, the median spin of actively growing BHs in our model does not reveal the presence of rapidly spinning BHs. This is demonstrated in the lower panel of Fig. 8 where we have selected the actively growing BHs in our sample ($\dot{m} > 0.01$) and plot the median of their spin distribution weighted by disc luminosity. As illustrated in the plot, the predictions are consistent with low spins at all redshifts even at $z \sim 0$. This is because in this sample we have selected only the actively growing BHs; we find that almost exclusively these are $\lesssim 10^9\,M_\odot$ BHs and thus have low spins. Hence, these BHs dominate the sample and determine the overall trend of the median.

The a–M_{BH} correlation can be further elucidated through the dependence of the accretion efficiency, ϵ, on the mass of actively growing BHs. This is illustrated in Fig. 10, where we show the median of the ϵ distribution for the sample of actively growing BHs (solid lines), again weighted by the disc luminosity. At $z = 0$, the median is approximately constant for $M_{BH} \lesssim 10^8$, with a typical value of ∼0.07. As implied by the 10–90 percentiles (dotted lines), the range of the typical ϵ values is very small and restricted to 0.06–0.08 ($a = 0.1$–0.5). For higher BH masses, the efficiency can reach significantly higher values. This is a manifestation of the fact that high-mass BHs have high spins and thus high accretion efficiencies when they accrete in the thin-disc regime. Fig. 10 also demonstrates that the dependence of the efficiency on the BH mass does not change with redshift. Hence, at all redshifts, BHs have very well determined accretion efficiencies. It is therefore only the accretion rate that regulates the luminosity output from an accreting BH.

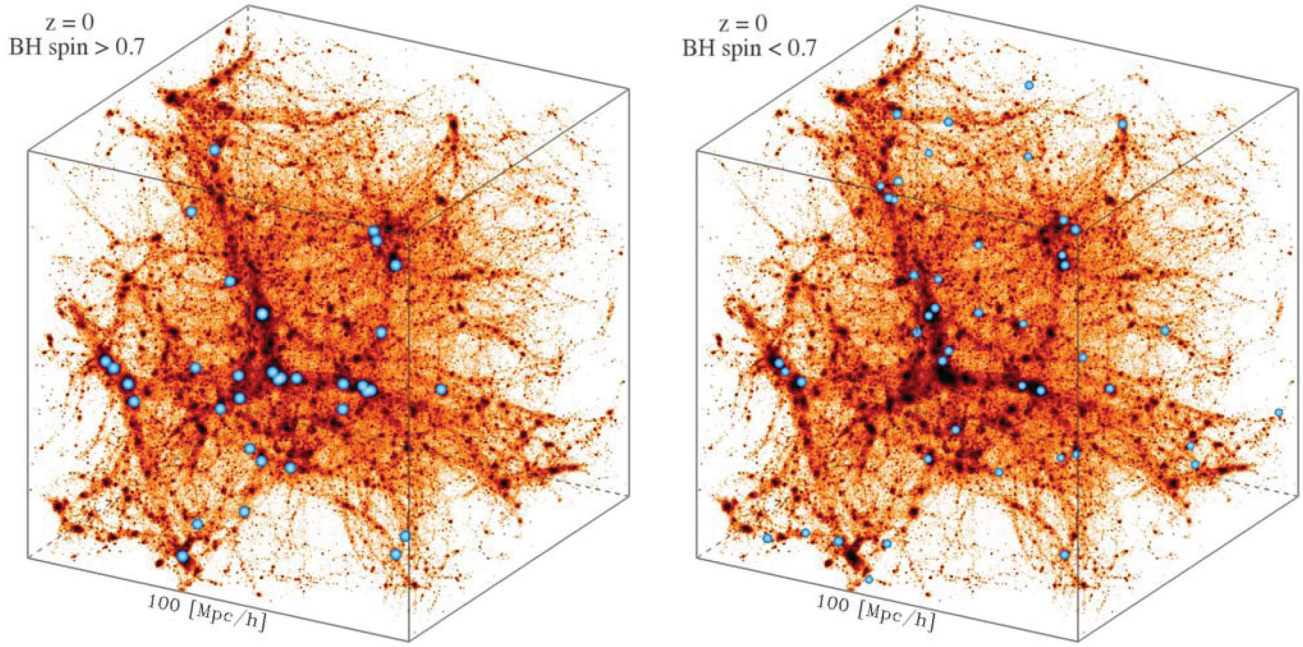

Figure 9. The distribution of rotating BHs at $z = 0$ in the Millennium simulation. The two panels show the same DM distribution in a cube of comoving length 100 Mpc h^{-1} colour-coded according to density (black represents the peaks of DM density). Overplotted with the blue spheres in the left-hand plot are galaxies with rapidly rotating BHs ($a > 0.7$) predicted by the semi-analytic model. Similarly, in the right-hand plot are shown a sample of galaxies with slowly rotating BHs ($a < 0.7$) randomly chosen from our data and equal in number to the $a > 0.7$ BHs. The size of the spheres is proportional to the spin of the BH that the galaxy hosts.

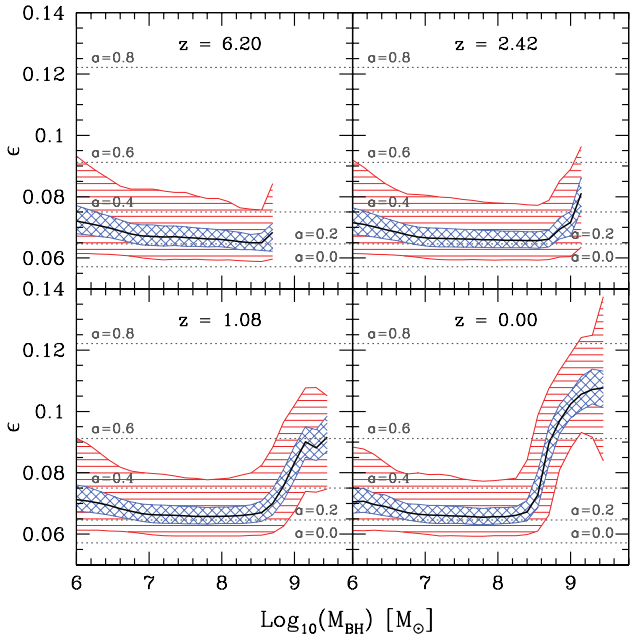

Figure 10. The median of the accretion efficiency, ϵ, as a function of BH mass at different redshifts (solid lines) for BHs that accrete in the thin-disc regime. Also shown are the 10–90 (red shaded areas) and 25–65 (blue shaded areas) percentile ranges of the ϵ–$M_{\rm BH}$ distribution. We also indicate the values of ϵ that correspond to the spin values of 0, 0.2, 0.4, 0.6 and 0.8 (dotted grey lines).

Note that BHs with given higher efficiencies ($\epsilon \gtrsim 0.15$) are still found in our BH populations. However, they are usually very massive ($M_{\rm BH} > 5 \times 10^8\,{\rm M}_\odot$) and do not undergo significant quasar activity. These BHs can support the formation of very strong jets in the presence of an ADAF and establish the host galaxy as a radio-loud AGN (high spins and BH mass are essential for high jet luminosities; Fanidakis et al. 2011). In this case, a plot similar to the one in the bottom panel of Fig. 8, weighted by jet instead of disc luminosity, will reveal a significant population of AGNs with rapidly rotating BHs. Hence, BH spins (and efficiencies) can display different distributions depending on the AGN population being probed.

4.3 The distribution of the Eddington ratio, $\lambda_{\rm Edd}$

Having explored the evolution of BH mass and spin, and calculated the accretion efficiencies for the BHs accreting in the thin-disc regime, we now investigate the disc luminosities of the accreting BHs predicted by the model. We calculate the bolometric disc luminosity in the ADAF and thin-disc regime using the formulation described in Section 4.1.

It is useful to scale $L_{\rm bol}$, in units of $L_{\rm Edd}$, in order to remove the dependence of the luminosity on the BH mass. For this purpose, we introduce the Eddington ratio, $\lambda_{\rm Edd} = L_{\rm bol}/L_{\rm Edd}$, defined as,

$$\lambda_{\rm Edd} = \begin{cases} 6.3 \times 10^{-5}\,\dot{m}, & \text{if } \dot{m} < 7.5 \times 10^{-6} \\ 44.4\,\dot{m}^2, & \text{if } 7.5 \times 10^{-6} < \dot{m} < 0.01 \\ \dot{m}, & \text{if } 0.01 \geq \dot{m} < \eta L_{\rm Edd} \\ \eta[1 + \ln(\dot{m}/\eta)], & \text{if } L_{\rm bol} \geq \eta L_{\rm Edd}. \end{cases} \quad (8)$$

In Fig. 11, we plot the distribution function of $\lambda_{\rm Edd}$ at different redshifts for all accreting sources with $M_{\rm BH} > 10^6\,{\rm M}_\odot$. The plane of the plot is divided into two distinct regions: the thin-disc regime and the ADAF regime (note the discontinuity due to the different dependence of $\lambda_{\rm Edd}$ on $\dot{m}$ in the ADAF and thin-disc regime). This distinction is essential since it will help us to unravel the space

© 2011 The Authors, MNRAS **419,** 2797–2820
Monthly Notices of the Royal Astronomical Society © 2011 RAS

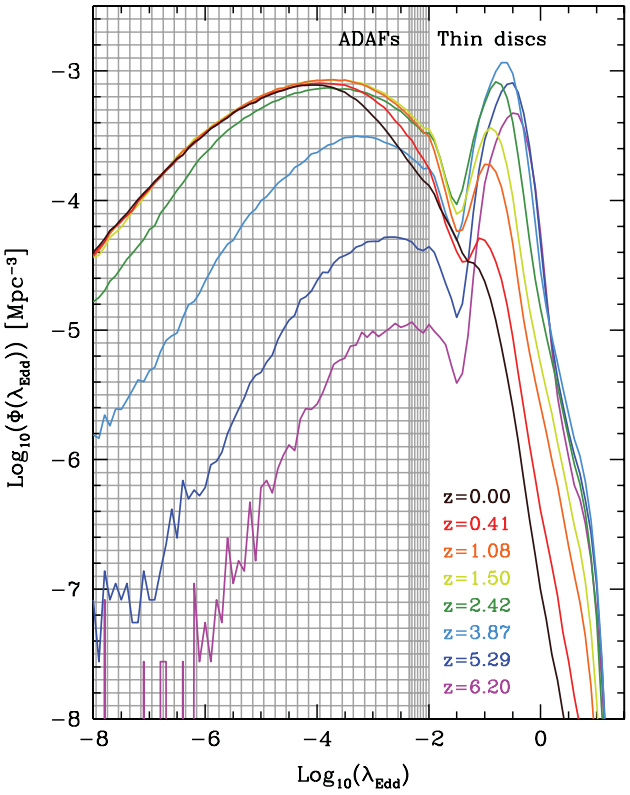

Figure 11. The distribution of $\lambda_{\rm Edd} = L_{\rm bol}/L_{\rm Edd}$ for different redshifts as indicated by the colour bar. The plane is divided into the radiatively-inefficient regime of ADAFs (shaded region) and radiatively-efficient regime of thin discs. The denser shading denotes the region of the plane where the discontinuous transition from the ADAF to the thin-disc regime takes place.

density and evolution of luminous and underluminous AGNs in our model.

The first important property that is clearly apparent in Fig. 11 is the bimodal nature of the $\lambda_{\rm Edd}$ distribution. The bimodality is exhibited at nearly all redshifts with both modes being characterized by a lognormal distribution. The first peak falls in the ADAF regime and it shifts from $\log_{10}\lambda_{\rm Edd} \simeq -2.5$ at $z \sim 6$ to $\log_{10}\lambda_{\rm Edd} \simeq -4$ at $z \sim 0$. The objects that contribute to that mode are all AGNs accreting during the hot-halo mode and also those AGNs accreting during the starburst mode with $\log_{10}\dot{m} < -2$. The second peak is located in the thin-disc regime and in this mode we find the most luminous objects in our model. Their peak shifts from $\log_{10}\lambda_{\rm Edd} \simeq -0.5$ at $z \sim 6$ to $\log_{10}\lambda_{\rm Edd} \simeq -1$ at $z \sim 0$, where it nearly disappears. The objects populating the second mode are exclusively AGNs accreting in the starburst mode. Since the thin-disc regime is radiatively efficient, these objects are expected to dominate the LF of AGNs in all bands (except the radio and perhaps the HXs).

The relative space density of the objects in each mode changes dramatically with redshift. At high redshifts, AGNs in the thin-disc regime are much more numerous than those in the ADAF regime. The BHs in these AGNs have low masses, accrete near the Eddington limit and double their mass several times within a few Gyr. In contrast, BHs in the ADAF regime are more than an order of magnitude less numerous. Their number density is dominated by BHs that have reached by that time a high mass (10^8–$10^9\,{\rm M}_\odot$) and their host galaxy undergoes a disc instability or minor merger event that provides gas at a low accretion rate. The first BHs accreting during the hot-halo mode also contribute to the objects populating the ADAF regime.

This picture changes in the low-redshift universe. As redshift decreases, more haloes enter the quasi-hydrostatic cooling regime and thus more BHs start to accrete in the ADAF regime (see also the evolution of the hot-halo accretion channel in Fig. 2). In this mode, we also find AGNs powered by gas accretion during gas-poor disc instabilities and galaxy mergers. Eventually, at $z = 0$, BHs accreting via an ADAF are much more common than those accreting via a thin disc.

Observational estimates of the $\lambda_{\rm Edd}$ distribution function based on optical emission lines also indicate the presence of two distinct modes of BH growth (Kauffmann & Heckman 2009). The first mode is associated with galactic spheroids that are undergoing significant SF and are characterized by a lognormal distribution of relatively high Eddington ratios. The second mode is linked to spheroids with low SF and governed by a power-law distribution of low Eddington ratios with a universal slope (at least in the range probed by observations). Our model suggests a similar behaviour since active galaxies in the thin-disc regime are starbursts, whereas in the ADAF regime, we have mostly passive galaxies with insignificant or no ongoing SF. In addition, $\lambda_{\rm Edd}$ clearly shows a lognormal distribution for starburst galaxies in our model. In a forthcoming paper, we will present the distribution of Eddington ratios for different galaxy populations, and BH masses, and explore the growth of BHs in these two regimes.

4.4 The $M_{\rm BH}$–$L_{\rm bol}$ correlation

When examining the distribution function of $\lambda_{\rm Edd}$ in Fig. 11, we find a small population of super-Eddington AGNs present at all redshifts. The number density of these AGNs drops sharply with increasing $\lambda_{\rm Edd}$ due to the logarithmic dependence on the accretion rate. Whether these objects represent the most luminous AGNs at some given redshift is not obvious. To unravel the relation between BH mass and accreted luminosity, we plot in Fig. 12 the median of the $L_{\rm bol}$–$M_{\rm BH}$ distribution and its associated percentiles at $z = 0.5$, 1 and 2. We show predictions both for objects accreting in the thin-disc (lower panels) and ADAF (upper panels) regimes. We distinguish with a different shading the region where $L_{\rm bol} \geq L_{\rm Edd}(M_{\rm BH})$.

When we consider BHs accreting in the ADAF regime, we find that the $M_{\rm BH}$–$L_{\rm bol}$ correlation is characterized by a floor at luminosities $\lesssim 10^{42}\,{\rm erg\,s^{-1}}$. In this regime, BHs have $\lambda_{\rm Edd} \simeq 10^{-4}$ and therefore they comprise the majority of accreting BHs in our model (remember the location of the peak of the $\lambda_{\rm Edd}$ distribution function in Fig. 11). For example, at $z = 0$, we find that ∼93 per cent of the BHs produce luminosities fainter than $10^{42}\,{\rm erg\,s^{-1}}$. For brighter luminosities, we find a strong correlation between $M_{\rm BH}$ and $L_{\rm bol}$ which indicates that the most massive BHs must have higher accretion rates compared to the lower mass ones. Indeed, when identifying the accretion properties of the most massive BHs, we find that the $\gtrsim 10^8\,{\rm M}_\odot$ BHs have $\lambda_{\rm Edd} \gtrsim 10^{-3}$. Thus, these BHs are able to produce a significant luminosity even though they accrete in the ADAF regime. In fact, we find $\sim 10^{10}\,{\rm M}_\odot$ BHs accreting at $\dot{m} \simeq 0.01$ which produce luminosities as high as $10^{46}\,{\rm erg\,s^{-1}}$. However, these BHs are very rare: at $z = 0.5$, where the space density of these BHs peaks, we find only a handful of them (space densities lower than $10^{-8}\,{\rm Mpc^{-3}}$). For higher redshifts, their space density declines, and as a consequence, the maximum disc luminosity produced in the ADAF regime is also reduced.

In the thin-disc regime, we find that accreting BHs typically produce luminosities greater than $10^{42}\,{\rm erg\,s^{-1}}$. The $M_{\rm BH}$–$L_{\rm bol}$

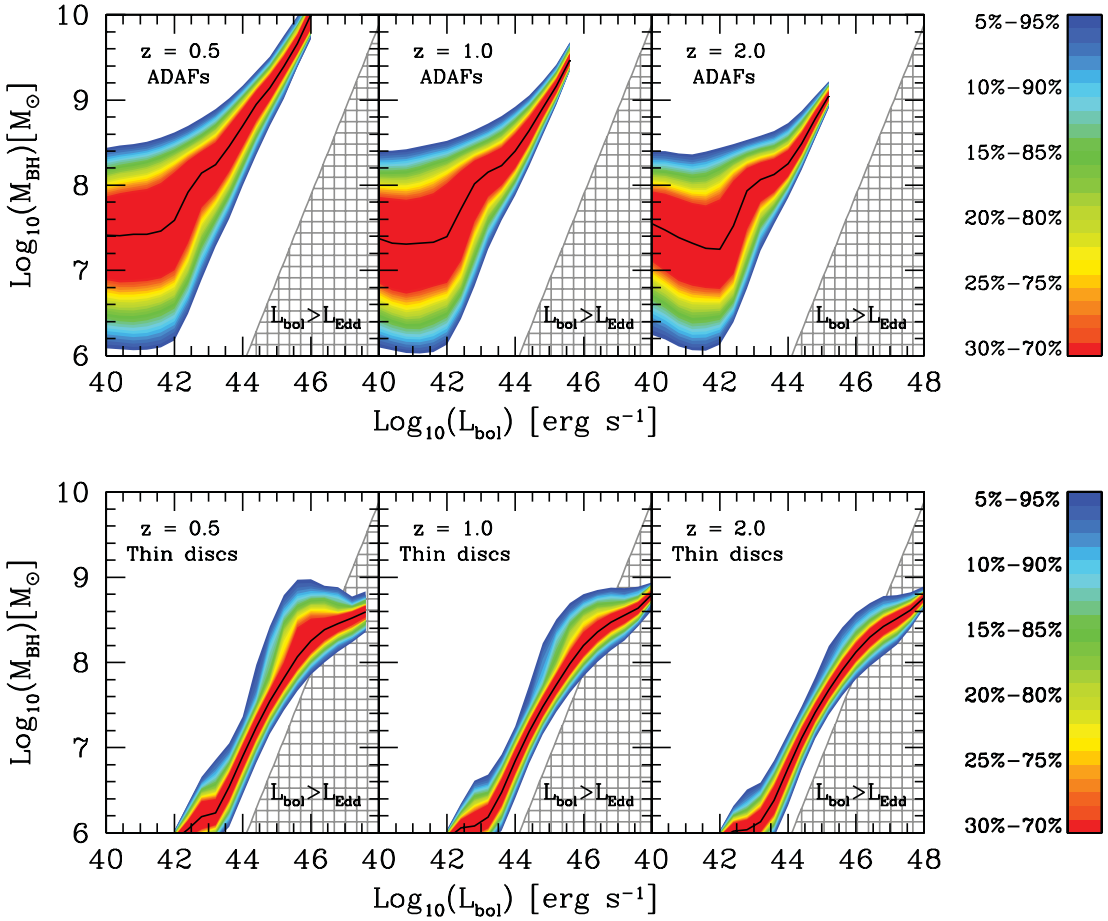

Figure 12. The median of the $M_{\rm BH}$–$L_{\rm bol}$ distribution at $z = 0.5$, 1 and 2 (black solid lines) for BHs accreting via a thin disc (lower panels) and ADAF (upper panels). The different percentiles of the distribution are colour-coded according to the bar on the right-hand side of each row. The shaded regions represent the regime where the accretion becomes super-Eddington.

correlation in this regime increases monotonically until the slope becomes significantly shallower near the highest luminosities achieved at a given redshift. The nature of the break in the slope is determined by the AGN feedback prescription in our model. When massive haloes reach quasi-hydrostatic equilibrium, they become subject to AGN feedback that suppresses the cooling flows; some of the mass which would have been involved in the cooling flow is instead accreted on to the BH. In these haloes, we find the most massive BHs in our model ($\gtrsim 10^9\,{\rm M}_\odot$, see Fig 7b). Therefore, these BHs are expected to live in gas-poor environments and when they accrete gas they usually do so via an ADAF disc. In this regime, the suppression of cooling flows forces the $M_{\rm BH}$–$L_{\rm bol}$ correlation to evolve only along the $L_{\rm bol}$ axis since accretion via a thin disc on to $\gtrsim 10^9\,{\rm M}_\odot$ BHs becomes very rare.

The correlation between $L_{\rm bol}$ and $M_{\rm BH}$ found at $z = 0.5$ remains approximately the same at higher redshifts. Only the ranges of BH mass corresponding to the bulk of the AGNs shift modestly to lower mass. This is due to the fact that in a hierarchical universe accretion shifts to lower BH masses at higher redshifts (Section 3). In addition, the break in the slope at high luminosities becomes less prominent since fewer haloes are in quasi-hydrostatic equilibrium at high redshifts.

The most luminous AGNs ($\gtrsim 10^{46}\,{\rm erg\,s^{-1}}$) are exclusively powered by super-Eddington accretion on to $\sim 10^8$–$10^9\,{\rm M}_\odot$ BHs. This implies that the most luminous quasars are expected to be found in $\sim 10^{12}$–$10^{13}\,{\rm M}_\odot$ DM haloes and not in the most massive ones (remember the $M_{\rm BH}$–$M_{\rm Halo}$ relation in Fig. 7b). This is in good agreement with the typical DM halo mass quasars inferred to inhabit as suggested by several observational clustering analyses (da Ângela et al. 2008; Ross et al. 2009; Shen et al. 2009). In contrast, accretion characterized by lower luminosities spans the whole range of BH masses (10^6–$10^{10}\,{\rm M}_\odot$).

5 THE EVOLUTION OF THE AGN LUMINOSITY FUNCTIONS

5.1 Bolometric corrections and obscuration

The LF is calculated in redshift ranges set by the observations against which we are comparing the model. The contribution of each AGN to the LF in a range $z = z_1$–z_2 is weighted by a factor

$$w_{\rm BH} = t_{\rm active}/\Delta t_{z_1,z_2}, \qquad (9)$$

where $\Delta t_{z_1,z_2}$ is the time-interval spanned by the redshift range z_1–z_2 and $t_{\rm active}$ is the time during which the AGN is 'on' during the $\Delta t_{z_1,z_2}$ interval (in principle, $t_{\rm active} = t_{\rm acc}$ only if the whole period of accretion falls within z_1–z_2). The bands for which we present predictions are the B band (4400 Å), SX (0.5–2 keV) and HX (2–10 keV). The bolometric corrections considered here are approximated by the following third-degree polynomial relations

© 2011 The Authors, MNRAS **419,** 2797–2820
Monthly Notices of the Royal Astronomical Society © 2011 RAS

(Marconi et al. 2004):

$$\log_{10}(L_{\rm HX}/L_{\rm bol}) = -1.65 - 0.22\mathcal{L} - 0.012\mathcal{L}^2 + 0.0015\mathcal{L}^3,$$
$$\log_{10}(L_{\rm SX}/L_{\rm bol}) = -1.54 - 0.24\mathcal{L} - 0.012\mathcal{L}^2 + 0.0015\mathcal{L}^3,$$
$$\log_{10}(\nu_{\rm B}L_{\nu_{\rm B}}/L_{\rm bol}) = -0.80 + 0.067\mathcal{L} - 0.017\mathcal{L}^2 + 0.0023\mathcal{L}^3, \quad (10)$$

where $\mathcal{L} = \log_{10}(L_{\rm bol}/{\rm L}_{\odot}) - 12$. Marconi et al. derive these corrections based on an observational spectral template where the spectrum at $\lambda > 1$ μm is truncated in order to remove the infrared (IR) bump. In this way, the spectrum corresponds to the intrinsic luminosity of the AGN (optical, UV and X-ray emission from the disc and hot corona). We apply equations (10) to both thin discs and ADAFs, though we note that there is evidence for a change in these corrections with $\lambda_{\rm Edd}$ (Vasudevan & Fabian 2007).

There is evidence that the fraction of obscured AGNs decreases with increasing X-ray luminosity (Steffen et al. 2003; Ueda et al. 2003; Hasinger 2004; La Franca et al. 2005), a trend found also by Simpson (2005) in a sample of broad and narrow-line AGNs from the SDSS. The question of whether the fraction of obscured AGNs depends also on redshift is more uncertain. If the obscuring medium is indeed gas and dust in the galaxy, then we would expect the fraction of obscured AGNs to be redshift-dependent (since in principle galaxy properties evolve with redshift). In addition, a strongly evolving population of obscured AGNs are required by AGN population synthesis models to reproduce the properties of the X-ray background (Comastri et al. 1995; Gilli, Risaliti & Salvati 1999; Ballantyne et al. 2006a; Ballantyne, Everett & Murray 2006b; Gilli, Comastri & Hasinger 2007). Ueda et al. (2003) and Steffen et al. (2003) suggest that such a trend is not clear in AGN samples from deep X-ray surveys. However, the analysis by Treister & Urry (2006) on AGN samples of higher optical spectroscopic completeness indicates that the relative fraction of obscured AGNs does increase with redshift.

More recently, Hasinger (2008) showed, based on a sample of X-ray-selected AGNs from 10 independent samples with high-redshift completeness, that the fraction of obscured AGNs increases strongly with decreasing luminosity and increasing redshift. According to Hasinger, the dependence of the obscured fraction of AGNs, $f_{\rm obsc}$, on $L_{\rm HX}$, can be approximated by a relation of the form

$$f_{\rm vis} = 1 - f_{\rm obsc} = 1 + 0.281\log_{10}\left(\frac{L_{\rm HX}}{10^{43.75}\,{\rm erg\,s^{-1}}}\right) - A(z), \quad (11)$$

where $A(z) = 0.308(1+z)^{0.48}$ and $L_{\rm HX} = 10^{42}$–10^{46} erg s^{-1}. A broken power law can also be used to describe $A(z)$. The best fit gives a power law of the form $A(z) \propto (1+z)^{0.62}$ that saturates at $z = 2.06$ and remains constant thereafter.

In principle, our model allows us to construct a physical prescription for the obscuration of the central engine, which could be coupled to the galactic disc geometry and its gas/dust content. However, to avoid adding an extra, uncertain, layer of complexity, we decided to use the empirical prescription of Hasinger (2008) to account for the effects of obscuration. To do so, we utilize the dependence of the obscured fraction on the $L_{\rm HX}$ luminosity found by Hasinger as follows. We calculate the fraction of visible AGNs, $f_{\rm vis} = 1 - f_{\rm obsc}$, at a given luminosity bin in the 2–10 keV band using equation (11) and then associate the value of $f_{\rm vis}$ with the corresponding B-band or SX luminosity bin using equations (10). The LF can then be expressed as

$$\left.\frac{{\rm d}\Phi}{{\rm d}\log_{10}(L_{\rm X})}\right|_{\rm vis} = f_{\rm vis}(L_{\rm HX}, z)\frac{{\rm d}\Phi}{{\rm d}\log_{10}(L_{\rm X})}. \quad (12)$$

In our analysis, we choose the single, instead of the broken, power-law form of $A(z)$ to allow for a comparison with previous work (e.g. La Franca et al. 2005). This provides a simple, yet well constrained, prescription for the effects of obscuration.

5.2 The optical LF

In Fig. 13, we present the $b_{\rm J}$-band LF of quasars in nine different redshift bins between $0.4 < z < 4.25$. Our predictions are shown before and after applying the obscuration prescription (dashed and solid black lines, respectively). Absolute magnitudes are first calculated in the B band using the standard expression

$$M_B = -10.44 - 2.5\log_{10}(\nu_B L_{\nu_B}/10^{40}\,{\rm erg\,s^{-1}}), \quad (13)$$

for magnitudes in the Vega system, and then converted to the $b_{\rm J}$ band in the AB system using the correction $M_{b_{\rm J}} = M_B - 0.072$ (Croom et al. 2009b). Predictions are shown for the F11b model and, for comparison, for the model (solid orange lines) presented in Fanidakis et al. (2011), that is, F11a. The F11a model was constrained to match the observed LF of quasars at $z \sim 0.5$. In brief, the model assumed a constant obscuration fraction of $f_{\rm vis} = 0.6$ and a $b_{\rm J}$-band bolometric correction of 0.2. Finally, our predictions for the $z = 0$ LF, after applying the obscuration effect, are shown in every panel (black dotted lines) to help assess the evolution with redshift.

Our predictions are compared to the 2SLAQ+SDSS QSO LFs. QSO magnitudes in the 2SLAQ sample were obtained in the g band and therefore need to be converted to the $b_{\rm J}$ band considered here. We use $M_{b_{\rm J}}(z=0) = M_g(z=2) + 0.455$ (Croom et al. 2009b), where the K-corrections for the g band have been normalized to $z = 2$. In addition, we plot the 2QZ LFs by Croom et al. (2004). The observed LFs agree very well with each other particularly at bright magnitudes ($M_{b_{\rm J}} < -24$). The modest disagreement seen at the faintest magnitudes is due to the different K-correction applied to the 2QZ QSO sample by Croom et al. (2004). For higher redshifts, we include the LF derived from the SDSS Data Release 3 sample (DR3, Richards et al. 2006). The SDSS DR3 LF is obtained in the i band; we convert it to the $b_{\rm J}$ band using $M_{b_{\rm J}}(z=0) = M_i(z=2) + 0.71$, assuming a spectral index of $a_\nu = -0.5$ (Croom et al. 2009b).

The AGN model contains one adjustable parameter that must be constrained by observational data. This is the $f_{\rm q}$ parameter in equation (2) which sets the accretion time-scale. We choose the $b_{\rm J}$ band for constraining the value of $f_{\rm q}$ because it is the most sensitive to the modelling of AGNs. Other free parameters are constrained by the galaxy formation model and are adjusted to reproduce the observed galaxy LFs (Bower et al. 2006; Lagos et al. 2011c). Since $f_{\rm q}$ is essentially the only free parameter in our model, the predictions presented in these sections are genuine predictions of the underlying galaxy formation model. By fixing $f_{\rm q}$ to 10, we obtain an excellent overall match to the observed QSO LFs in the $0.4 < z < 2.6$ redshift interval. For higher redshifts, our model provides a reasonable match, but it predicts a steeper slope for the bright end compared to the SDSS DR3 LF.

A comparison between the predictions for the $z = 0.4$–4.25 and $z = 0$ LFs in Fig. 13 shows that quasars undergo significant cosmic evolution. For example, quasars with $M_{b_{\rm J}} = -25$ increase in space density from $\sim 10^{-8}$ to $\sim 10^{-6}$ Mpc^{-3} between $z = 0$ and $z = 2.2$–2.6. This strong evolution in the space density of quasars is due to the fact that disc instabilities and galaxy mergers, the two processes that trigger accretion during the starburst mode, become more frequent at higher redshifts (see Fig. 1). Yet, the strong evolution of quasars does not affect only their space density. At a fixed space density

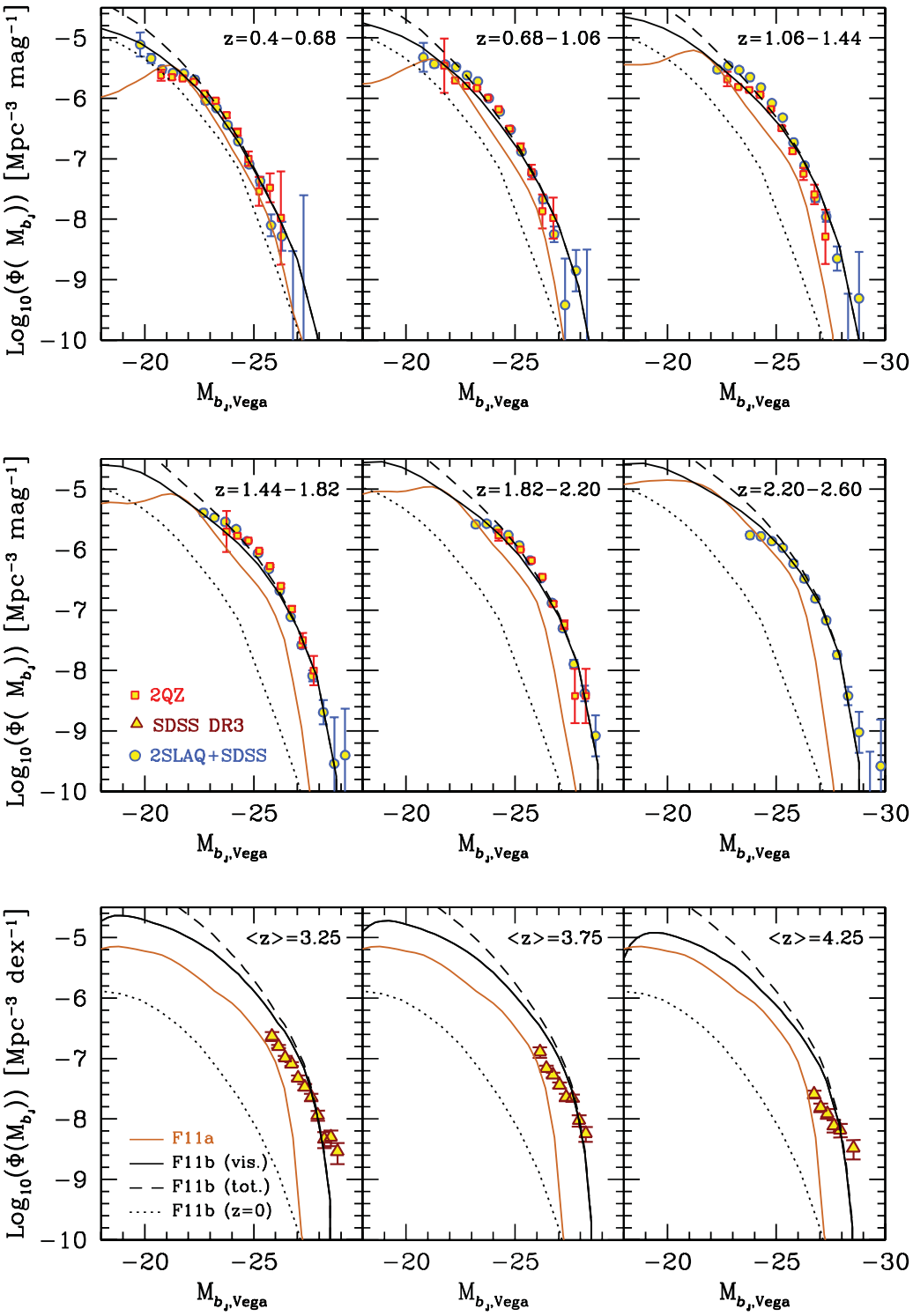

Figure 13. The b_J-band quasar LF in the redshift range $0.4 < z < 4.25$. Predictions are shown for the model described in this paper (F11b; solid black lines) and the F11a (solid orange lines) model. For reference, we show the $z = 0$ LF in the F11b model in every panel (dotted black line). We also show the evolution of the LF in the F11b model before applying the obscuration (dashed black lines). The observations correspond to the LFs estimated from the 2SLAQ+SDSS (blue-filled circles, Croom et al. 2009b), 2QZ (red-filled circles, Croom et al. 2004) and SDSS (brown-filled triangles, Richards et al. 2006) surveys.

of 10^{-8} Mpc^{-3}, the quasar LF brightens from $M_{b_J} \simeq -25$ at $z = 0$ to $M_{b_J} \simeq -28$ at $z = 2.2$–2.6. This is attributed to the fact that the cold gas becomes more abundant with increasing redshift. This is equivalent to a strong increase in the gas reservoir available for feeding the central BHs in a starburst.

The processes that are responsible for driving the formation of stars in starbursts are also responsible for the cosmic evolution of quasars. Qualitatively, the strong link between the formation of stars and quasar activity (and therefore BH growth) can be illustrated by comparing the F11a and F11b models. In Fig. 1, we can see

© 2011 The Authors, MNRAS **419,** 2797–2820
Monthly Notices of the Royal Astronomical Society © 2011 RAS

Table 4. Typical values of $f_{\rm vis}$ in three $b_{\rm J}$ magnitude bins (and corresponding HX luminosities in units of erg s^{-1}) and its evolution with redshift.

	$M_{b_{\rm J}} = -20$ ($L_{\rm HX} \simeq$s $10^{43.3}$)	$M_{b_{\rm J}} = -24$ ($L_{\rm HX} \simeq 10^{44.5}$)	$M_{b_{\rm J}} = -28$ ($L_{\rm HX} \simeq 10^{45.6}$)
$z = 0.4$	53.3 per cent	84.3 per cent	100 per cent
$z = 0.68$	49.0 per cent	81.0 per cent	100 per cent
$z = 1.06$	44.9 per cent	76.9 per cent	100 per cent
$z = 1.44$	42.2 per cent	73.3 per cent	100 per cent
$z = 1.82$	27.8 per cent	69.9 per cent	100 per cent
$z = 2.20$	24.7 per cent	66.7 per cent	98.7 per cent

that the SFR density in bursts in the F11a model shows a milder evolution with redshift compared to the F11b model. Hence, the model predicts less evolution in the amount of gas available for accretion which then results in more modest evolution of the quasar LF. As a consequence, the predictions of the F11a model provide an overall poor match to the observed LFs.

The evolution of the SFR density with redshift does not imply a proportional evolution of quasar luminosities in the $M_{b_{\rm J}}$–$\Phi(M_{b_{\rm J}})$ plane. The slope of the bright end of the LF becomes progressively steeper with increasing redshift, whereas the slope of the faint end remains almost unchanged. This differential evolution is determined exclusively by the accretion physics in our model. In the $0.01 < \dot{m} < 1$ regime, the disc luminosity scales in proportion to $\dot{m}$. However, when the flow becomes substantially super-Eddington, the luminosity instead grows as $\ln(\dot{m})$ and therefore the range of predicted luminosities decreases dramatically. The logarithmic dependence of luminosity on the accretion rate has a strong impact on the shape of the LF, resulting in a very steep slope at the bright end. This becomes apparent in the highest redshift intervals ($z > 1.44$), where the relative number of super-Eddington accreting sources becomes significantly higher than in the lower redshift intervals.

Another factor that influences the evolution of the quasar LF is obscuration. Low-luminosity quasars are heavily obscured according to our obscuration prescription, and thus remain well buried in their host galaxies. This is clear from Table 4 where we summarize the evolution of the $f_{\rm vis}$ value in the redshift range of interest. In the highest redshift intervals, the obscuration becomes more prominent affecting even the brightest sources. However, the intense accretion activity during the starburst mode dominates the evolution of the faint end and therefore a strong cosmic evolution in the space density of the faintest sources, similar to that of the brightest sources, is still observed. None the less, the obscuration significantly influences the cosmic abundance of the quasar populations dominating the faint end, something that will become more evident in Section 5.5 where we will explore the cosmic evolution of quasars of given intrinsic luminosity. We note that if we assume the recipe of constant obscuration from Fanidakis et al. (2011), namely $f_{\rm vis} = 0.6$, we find a space density for the $M_{b_{\rm J}} > -20$ quasars that is 0.15 dex higher than the one predicted by the Hasinger (2008) recipe. For brighter quasars, assuming such a recipe makes a negligible difference.

Finally, we point out that those AGNs that are powered by an ADAF contribute significantly to the space density of AGNs fainter than $M_{b_{\rm J}} \simeq -22$ in the low-redshift universe. In fact, in the range $0.4 < z < 0.68$, we find ADAF sources with $\dot{m} \simeq 0.001$–0.01 contributing also to the knee of the LF. These systems represent the rare $\gtrsim 10^9$ M$_\odot$ accreting BHs in Fig. 12. Identifying the ADAF systems with optically-bright quasars introduces a caveat for the model since quasars typically show high-excitation spectra which indicate the presence of a bright, UV thin disc (see e.g. Marchesini, Celotti & Ferrarese 2004). None the less, these systems have accretion rates in the regime where the transition to a thin disc takes place. Observations of stellar mass BH binary systems show that this transition is complex, probably taking on a composite structure with the thin disc replacing the hot flow at progressively smaller radii (see e.g. the review by Done et al. 2007). Such a configuration could possibly produce high-excitation lines while retaining also the ADAF characteristics. For a more comprehensible representation of the objects populating the LF, we refer the reader to Fig. 16 in Section 5.4 where we decompose the LF into the contribution from ADAF, thin-disc, and also super-Eddington sources.

5.3 The X-ray LFs

The good agreement between our model predictions for the optical LF and observations motivates us to study also the evolution of X-ray AGNs as a function of redshift and intrinsic luminosity. X-rays account for a considerable fraction of the bolometric luminosity of the accretion disc, in fact more than the optical in low-luminosity AGNs, and therefore provide an ideal band for detecting AGNs in a broader luminosity range. Hence, AGNs selected in X-rays are a more representative sample of the AGN population at some given redshift. The observed continuum spectra of AGNs in the X-rays can be represented by a simple power law in the *hard* part of the spectrum (2–10 keV), with a universal spectral index of $\alpha_{\rm X} \sim 0.7$ which is independent of the luminosity over several decades. In the *soft* part of the spectrum (0.5–2 keV), the power-law continuum is strongly attenuated by absorption, which is typically in excess of the Galactic value, and corresponds to a medium with hydrogen column densities of $N_{\rm H} \simeq 10^{20}$–10^{23} cm^{-1} (Reynolds 1997; George et al. 1998; Piconcelli et al. 2005).

The cosmic evolution of X-ray AGNs has been investigated by employing AGNs selected both in the soft (Miyaji et al. 2000; Hasinger et al. 2005) and in the hard part of the X-ray spectrum (Ueda et al. 2003; Barger et al. 2005; La Franca et al. 2005; Aird et al. 2010). Mirroring the strong evolution of the optical LF, the LF of X-ray-selected AGNs also evolves strongly with cosmic time. The LFs in SXs and HXs have similar characteristics, evolving at the same rate, but differ significantly in normalization. For example, the Hasinger et al. (2005) SX sample has approximately a five times smaller global LF normalization compared to that of the HX LF estimated by Ueda et al. (2003). This difference is most likely attributed to the fact that a large fraction of AGNs are obscured in the SXs (type-2 AGNs are by a factor of 4 more numerous than type-1 AGNs, Risaliti, Maiolino & Salvati 1999).

We present in this section our predictions for the X-ray LFs. We calculate the X-ray emission in the 0.5–2 and 2–10 keV bands using equation (10). Our predictions for the SX and HX LFs are shown in Figs 14 and 15, respectively (solid lines). The predictions are compared to the LFs estimated by Hasinger et al. (2005) for the SXs and Ueda et al. (2003) for the HXs. We do not show estimates from observations that consist only of upper limits (empty bins). The Hasinger et al. sample comprises unobscured AGNs and therefore we need to take the effect of obscuration into account for the sample in the SXs. We do so by using the Hasinger (2008) prescription as explained in the previous section. We note that the obscuration fraction in the AGN sample from Hasinger (2008) is measured only for luminosities $10^{42} < L_{\rm HX} < 10^{46}$ erg s^{-1}. For luminosities lower than 10^{42} erg s^{-1}, we assume that $f_{\rm vis}$ remains constant and equal to $f_{\rm vis}(10^{42}$ erg s$^{-1})$. To quantify the effect of absorption, we also plot in the SXs the total population of AGNs (dashed lines).

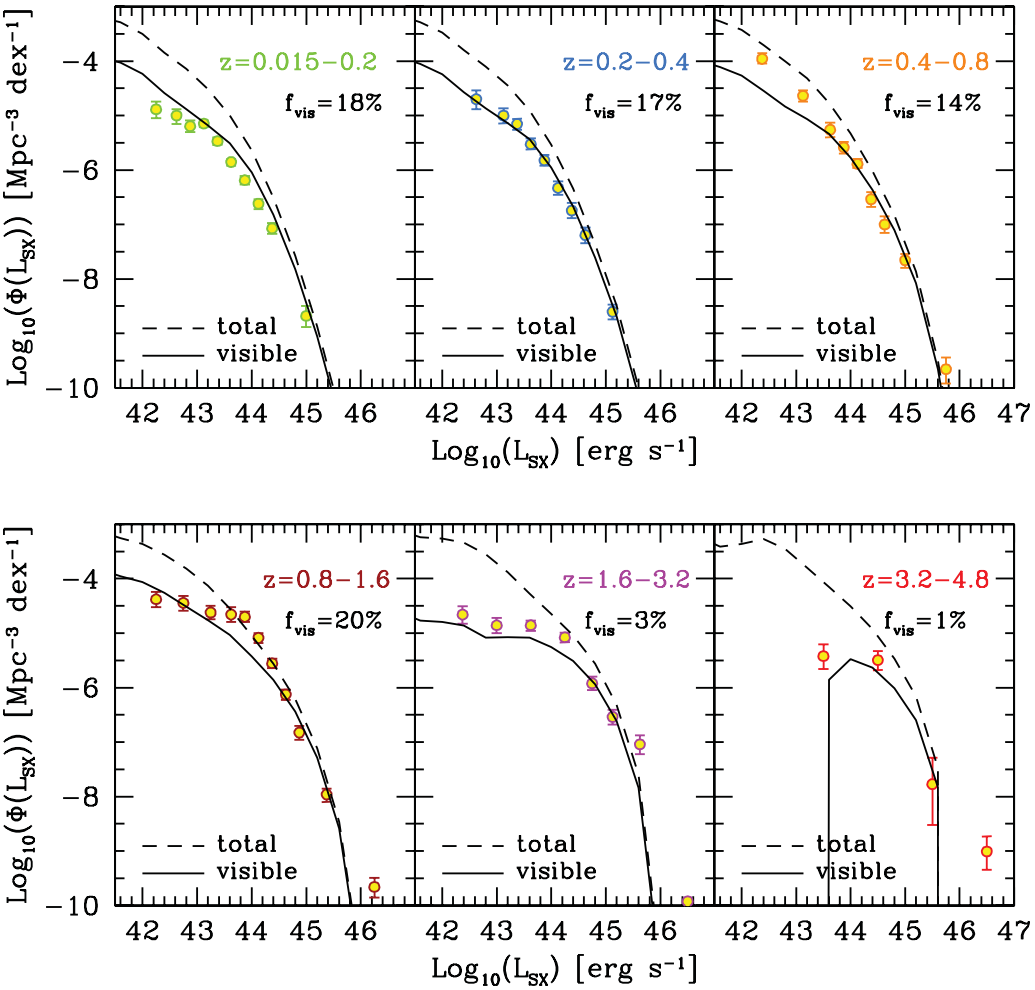

Figure 14. The predictions of our model for the evolution of the SX LF (solid black lines). Predictions are shown before (total: dashed lines) and after (visible: solid lines) we apply the effects of obscuration. In addition, we quote the total fraction of AGNs with $L_{\rm SX} > 10^{42}\,{\rm erg\,s^{-1}}$ that are visible in every luminosity bin. Data are taken from Ueda et al. (2003).

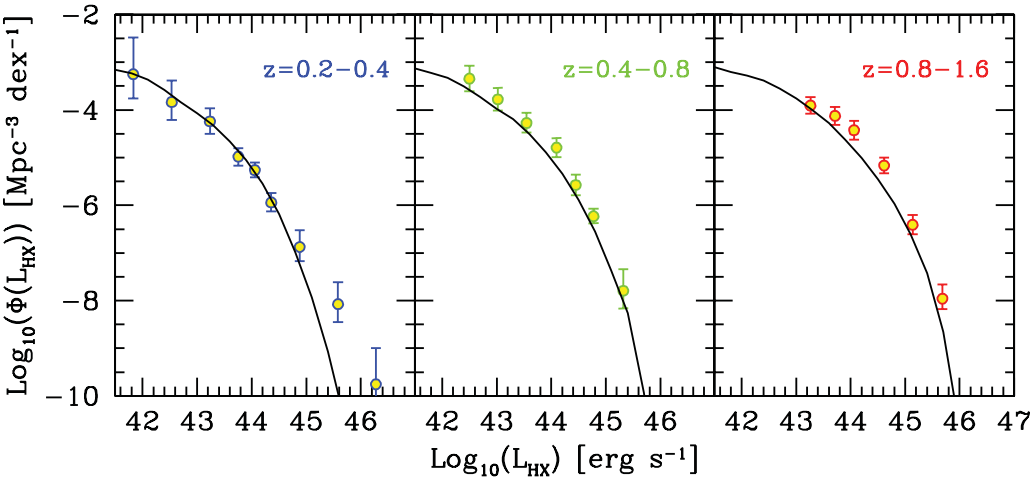

Figure 15. The evolution of the HX LF in our model (solid black lines). Predictions are shown only for $z < 1.6$ and no obscuration correction is applied. Observational data are taken from Hasinger et al. (2005).

Overall, our model provides a very good match to the observations in both SXs and HXs. We find discrepancies between the observations and the model predictions in the redshift range $z = 0.4 - 0.8$, where the space density of the $<10^{43}\,{\rm erg\,s^{-1}}$ SX AGNs is underpredicted. A similar disparity is also seen in the $z = 0.8$–1.6 bin for the 10^{43}–$10^{45}\,{\rm erg\,s^{-1}}$ AGNs. However, it is important to bear in mind that our predictions in the SXs depend strongly on our prescription for the obscuration. Since our predictions for the

© 2011 The Authors, MNRAS **419,** 2797–2820
Monthly Notices of the Royal Astronomical Society © 2011 RAS

total AGN population are always well above the observations, it is likely that these discrepancies can be attributed to the modelling of the obscuration. We also find that the model predicts at high redshifts a somewhat steeper bright end compared to the observations, although the very brightest point could be due to beaming effects in a small fraction of the more numerous lower luminosity sources (Ghisellini et al. 2010). None the less, the very steep slope in our predictions is again a manifestation of the Eddington limit applied to the super-Eddington sources, an effect that becomes more significant at higher redshifts as already explained in the previous section.

As illustrated by the predictions for the SX LF in Fig. 14, the space density of the faintest SX AGN in our model is strongly affected by obscuration. As illustrated by the LFs in the SXs, AGNs with $L_{\rm SX} \simeq 10^{42}$–$10^{44}\,{\rm erg\,s^{-1}}$ are heavily obscured. In fact, in the redshift interval $z = 0.015$–0.2, only 19 per cent of the total AGN population with $L_{\rm SX} \geq 10^{42}\,{\rm erg\,s^{-1}}$ is visible in SXs. This fraction of visible AGNs becomes even smaller at higher redshift because the obscuration becomes more prominent with redshift. At $z > 3.2$, only a negligible fraction of the AGNs are seen in SXs. Based on these results, we estimate that approximately 10 per cent of the total number of AGNs with $L_{\rm SX} \geq 10^{42}\,{\rm erg\,s^{-1}}$ are visible in the SXs in the $z = 0.015$–4.8 universe.

5.4 The bolometric LF

We have so far studied the LF of AGNs and quasars in different wavebands using the bolometric corrections in equation (10) and compared them to the available observations. We find that our predictions match reasonably well the observations, particularly in the $b_{\rm J}$ band. Further insight into the evolution of AGNs can be gained by calculating the bolometric LF and studying its evolution with redshift. This will provide an overall characterization of the global AGN population and might reveal additional trends which are perhaps unseen when exploring the LF in the other bands we studied earlier.

To calculate the bolometric LF, we consider all accreting objects (in both the starburst and hot-halo mode) taking into account the regime in which these objects accrete (ADAF or thin disc). This time we probe a wider baseline in luminosity and thus we expect to see clearly the contribution of the AGN powered by ADAFs. We calculate the bolometric LF at some redshift z following the technique described in the previous sections. AGNs are sampled over a period equal to 20 per cent of the age of the universe at redshift z. We need to stress that the resolution of the Millennium simulation has an impact on the gas properties of the $\lesssim 10^{44}\,{\rm erg\,s^{-1}}$ objects at high redshifts. Therefore, we compare the N-body results with high-resolution simulations using Monte Carlo (MC) halo merger trees to test our predictions. The MC algorithm we use to generate the DM halo merger trees has been presented in Parkinson, Cole & Helly (2008). The algorithm is a modification of the Extended Press–Schechter algorithm described in Cole et al. (2000) and accurately reproduces the conditional MFs predicted by the Millennium N-body simulation. The comparison shows that the calculation with N-body trees becomes incomplete for $L_{\rm bol} \lesssim 10^{44}$–$10^{45}\,{\rm erg\,s^{-1}}$ AGNs at $z \gtrsim 4$. Therefore, at high redshifts, we show predictions only down to the resolution limit of the Millennium simulation (the limit is indicated by the vertical dotted lines).

Our predictions for the bolometric LF are shown in Fig. 16 (solid black lines) in the redshift range $z = 0.2$–6. In addition, we show independently the contribution to the LF from the ADAF, thin-disc, and $L_{\rm bol} \geq L_{\rm Edd}$ sources (dashed coloured lines). Our results are compared to the bolometric LF estimated from observations by Hopkins et al. (2007), using a combination of a large set of observed LFs in the optical, X-ray, near- and mid-IR wavebands, and model-dependent corrections. The authors are able to reproduce the bolometric LF and the individual bands by employing their best-fitting estimates of the column density and spectral energy distribution, but also a prescription for the obscuration fraction which is assumed to be a function of the luminosity (though not of redshift). The data used by Hopkins et al. cover wide redshift ranges and sometimes the same data sets appear in more than one bin (especially the SX data at $z > 0.8$). Therefore, comparison between the data and our predictions should be treated with caution because the observations over such wide redshift bins could hide evolutionary trends.

Nevertheless, our predictions for the bolometric LF match very well the observational estimates across a wide range of redshifts. In the low-redshift universe ($z \lesssim 1$), the model reproduces the faint and bright ends of the observed LF remarkably well. At the faint end, the model predicts a significant contribution from AGNs powered by ADAFs. These are predominately massive BHs accreting at low accretion rates in the hot-halo mode and they account for the increasing abundance with decreasing redshift of the faint AGN (see also Fig. 11). By contrast, the bright end is always populated by AGNs radiating near or at greater than the Eddington limit. These are the AGNs that harbour the rapidly growing 10^7–$10^8\,{\rm M_\odot}$ BHs in the low-redshift universe.

At high redshifts ($z \gtrsim 1$), the contribution from the ADAF sources becomes less important. This is not only because fewer haloes reach quasi-hydrostatic equilibrium, but also because most of the accretion during the starburst mode takes place in the thin-disc regime. The LF at these redshifts is dominated by AGNs whose BHs accrete during the starburst model. The BHs in these AGNs grow very fast and they usually double their mass several times during a single accretion episode. This gives rise to a strongly evolving population of super-Eddington sources that dominate the bright end of the LF.

The bright end of the model $z \geq 1$ bolometric LF tends to have a steeper slope than the observational estimate, indicating a deficiency of very bright quasars ($L_{\rm bol} > 10^{48}\,{\rm erg\,s^{-1}}$) in the model.

This merely reflects the constraint on super-Eddington luminosities in the model. As discussed before, this suppression (equation 8) results in a very steep decline in the space density of super-Eddington sources, an effect clearly seen earlier in the optical LF. This is illustrated in Fig. 17(a), where we show how the LF at $z = 2$ depends on the parameter η of equation (8) which controls the suppression of super-Eddington sources. Above an accretion rate $\dot{m} = \eta$, the bolometric luminosity is assumed to increase only logarithmically with $\dot{m}$ as $\ln(\dot{m}/\eta)+1$. Interestingly, when no Eddington limit ($\eta \gg 1$) is imposed, the model can account for the brightest quasars observed. It is important to note, however, that many of the brightest luminosity bins do not include any observed objects and are therefore upper limits.

To allow sufficiently high quasar luminosities, we need to allow accretion rates in the most massive BHs much above the Eddington rate.[3] This is necessary because even if we assume that all the BHs accrete at the Eddington limit, the maximum luminosity that results is $\sim 10^{48}\,{\rm erg\,s^{-1}}$ (when we consider, for example, a 10^{10}-$\,{\rm M_\odot}$ BH) still lower than the highest luminosities observed. Note, however,

[3] For comparison, a BH accreting at $\dot{m} \sim 100$ is only approximately five times more luminous than a BH accreting at $\dot{m} = 1$.

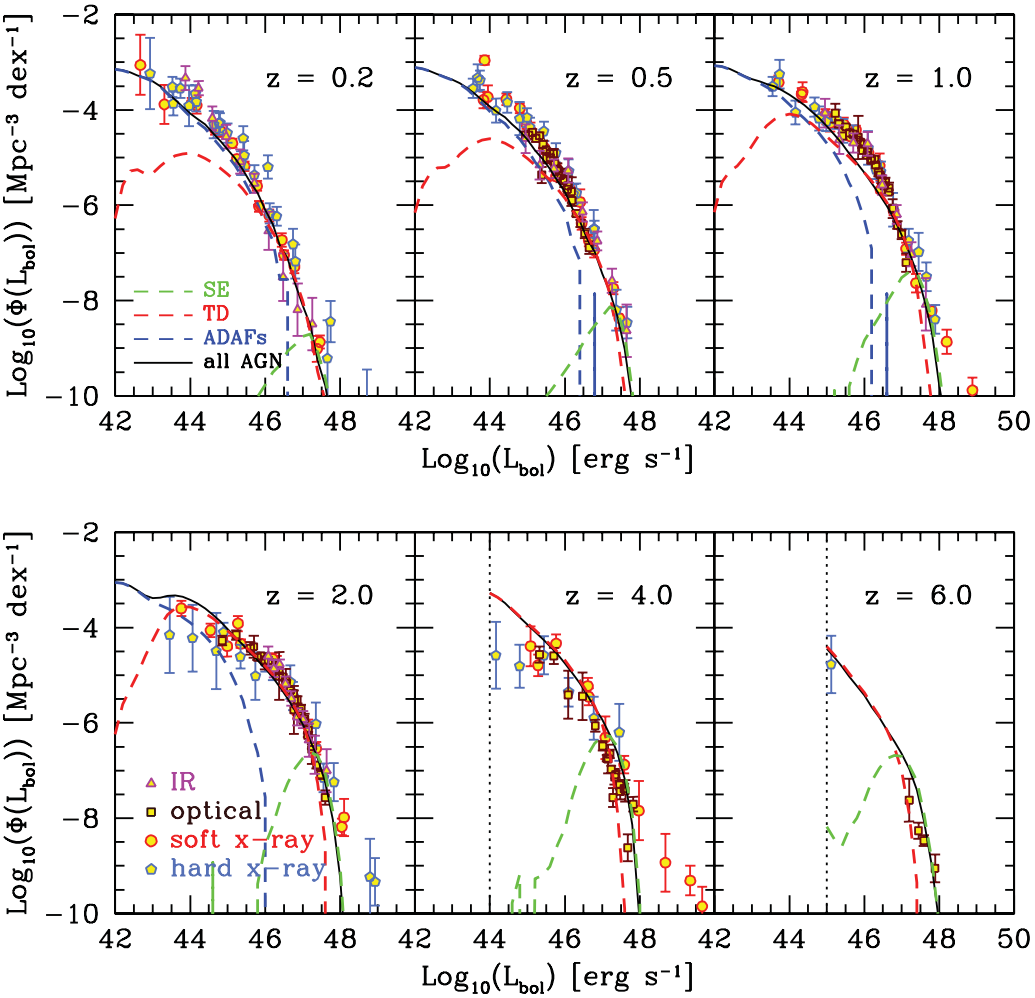

Figure 16. The bolometric LF predicted by our model in six redshift bins between $z = 0.2$ and 6. Also shown is the bolometric LF estimated by Hopkins, Richards & Hernquist (2007) using a large sample of IR, optical, SX and HX LFs (we refer the reader to Hopkins et al. 2007 for a description of the observational samples). The bolometric LFs inferred by the different bands are depicted by different colours as indicated by the labels. The dashed lines show the contribution to the LF of the ADAF (blue), thin disc (red) and super-Eddington sources (green). The vertical dotted lines indicate the resolution limit of the Millennium simulation.

that AGN feedback prevents BHs more massive than $10^9\,M_\odot$ from accreting at or near supper-Eddington accretion rates in our model.

To allow larger accretion rates on to BHs, we can either increase the amount of gas that is fed into it or decrease the accretion time-scale. The former solution corresponds to increasing the value of $f_{\rm BH}$ (the parameter that determines the fraction of gas available for accretion) which has already been tuned to provide a good match to the local BH density and MF. The latter solution gives us the freedom to adjust the accretion rates without changing the BH mass properties. By decreasing the value of $f_{\rm q}$ in equation (2), we can obtain higher $\dot{m}$ values and therefore boost the resulting bolometric luminosities to $L_{\rm bol} > 10^{48}\,{\rm erg\,s^{-1}}$. This is illustrated in Fig. 17(b), where we show the bolometric LF at $z = 2$ assuming $f_{\rm q} = 20$, 10, 1 and 0.5. However, changing the value of $f_{\rm q}$ has a strong effect on the faint end of the LF since it decreases the space density of the low-accretion-rate objects (the space density of the $\lesssim 10^{43}\,{\rm erg\,s^{-1}}$ AGN remains unchanged since that part of the LF is dominated by AGNs in the hot-halo mode in which the accretion time-scale is derived directly from the cooling time-scale of the gas). In addition, $f_{\rm q} \leq 1$ reduces significantly the duty cycle of actively growing BHs by limiting the typical accretion time-scale to $\sim 10^5$–10^6 yr. We conclude that having a constant value of $f_{\rm q}$ for all the accreting BHs limits our ability to account for the entire $L_{\rm bol}$ baseline. This indicates that a more sophisticated treatment of the accretion time-scale is therefore needed. However, this is beyond the scope of the present analysis. Again, we stress that sources in the brightest bins of the LF might be beamed and therefore represent much fainter systems with enhanced brightness (see Plotkin et al. 2010, for beamed quasars in the SDSS DR7 sample).

5.5 The evolution of cosmic AGN abundances: cosmic downsizing?

Our predictions for the optical, X-ray and bolometric LFs suggest that AGNs undergo important cosmic evolution. Their space density was significantly higher at earlier epochs, an evolutionary trend which suggests that AGN activity in the past was much more intense. Interestingly, our predictions indicate that the effects of the different physical processes considered here (the two accretion modes and the obscuration prescription) must have an important, and indeed variable, effect on the evolution of faint and bright AGNs. Therefore, to gain more insight into the cosmic evolution of AGNs, we study how

© 2011 The Authors, MNRAS **419,** 2797–2820
Monthly Notices of the Royal Astronomical Society © 2011 RAS

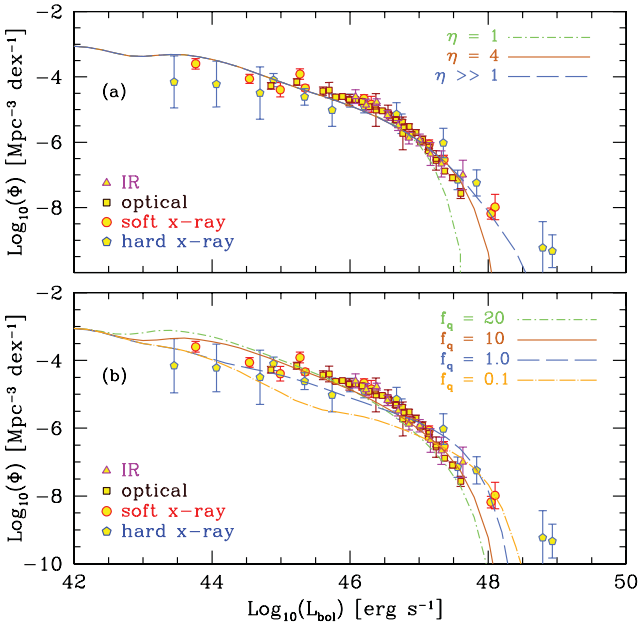

Figure 17. Panel (a): the dependence of the model bolometric LF on the assumptions for suppressing super-Eddington luminosities as parametrized by η in equation (8). Results are shown for the bolometric LF at $z = 0$ and for the three different luminosity limits indicated by the legends. Data are taken from Hopkins et al. (2007). Panel (b): our predictions for the bolometric LF at $z = 2$ for four different values of the accretion time-scale parameter, f_q, as indicated in the legend. Data are taken from Hopkins et al. (2007).

the abundance of different luminosity AGN populations evolves with redshift. This is shown in Fig. 18, where we show the cosmic evolution of optically selected AGNs in 11 absolute magnitude bins (left-hand panel) and SX-selected AGNs in eight luminosity bins (right-hand panel). We have shown separately the evolution of the total number of AGNs (before applying obscuration; dashed lines) and the visible population of AGNs (after applying obscuration; solid lines) in each bin. In this figure, we also show the sum over all magnitudes, $M_{b_J} \leq -18$, and luminosities, $\log_{10} L_{SX} \geq 10^{42}$ erg s^{-1} (black solid lines for the visible and black dashed lines for the total populations).

A feature immediately evident in both wavebands is that the evolution of the space density of AGNs is shallower for fainter sources. For example, in the optical waveband, the increase in the space density of the $M_{b_J} = -21$ AGN can be characterized in the redshift interval $z = 0$–1 by a power law of the form $(1 + z)^{\beta}$ with a very shallow slope of $\beta \sim 0.5$. The slope becomes gradually steeper reaching a value of $\sim$2.3 for the $M_{b_J} = -26$ AGN. The steepening of the slope is evident in both the total and the visible populations and it is driven primarily by the two distinct accretion channels in our model. The faintest sources depicted in the plots of Fig. 18 comprise a combination of objects accreting during the starburst and hot-halo modes. The hot-halo mode is responsible for building up the majority of them in the low-redshift universe, whereas the starburst mode accounts for them in the high-redshift universe (cf. Figs 11 and 12, but also Fig. 16). The slope characterizing the space density of these populations changes very slowly since the transition from the hot-halo-mode to starburst-mode domination is

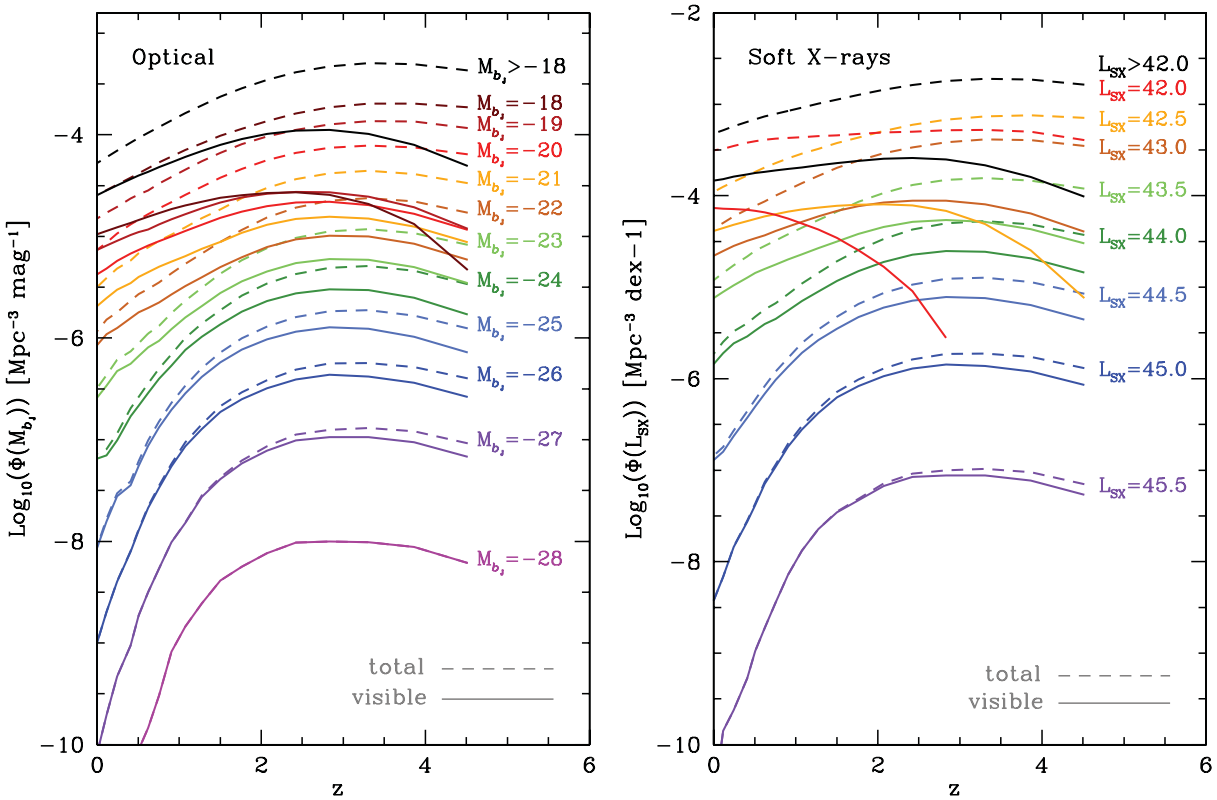

Figure 18. The cosmic evolution of different magnitude and luminosity classes of AGNs selected in the optical M_{b_J} band (left-hand panel) and SXs (right-hand panel), respectively. Also shown is the cosmic evolution of the sum of all AGNs with $M_{b_J} \leq -18$ in the optical and $\log_{10} L_{SX} \geq 10^{42}$ erg s^{-1} in the SXs (black lines). In all cases, we show with the dashed and solid lines the evolution of the space density of AGNs before (total) and after (visible) applying the effects of obscuration.

© 2011 The Authors, MNRAS **419,** 2797–2820
Monthly Notices of the Royal Astronomical Society © 2011 RAS

not characterized by a strong change in the space density of AGNs. However, when we consider brighter populations, the contribution of the hot-halo mode becomes less important (or vanishes for the brightest populations). In the low-redshift universe, we find a much steeper slope for the bright AGN because their evolution is driven primarily by the starburst mode, and starbursts are prevalent at high redshifts.

Trends similar to those seen in the cosmic evolution of AGNs in our model have been reported in the literature based on observational data. For example, Hopkins et al. (2007) find in their analysis of the bolometric LF that faint AGNs show a much more modest evolution than bright AGNs (fig. 9 in their paper; see also Ueda et al. 2003; Hasinger et al. 2005; La Franca et al. 2005; Croom et al. 2009b). These trends pose the following key question. Why does AGN activity shift from high-luminosity objects in the high-redshift universe to low-luminosity objects in the low-redshift Universe? Is this downsizing clear evidence against hierarchical galaxy formation models as has been interpreted by many authors? In our model, the downsizing of AGNs is manifested in the shallow slope that characterizes the evolution of faint AGNs. This arises naturally from the interplay between the starburst and hot-halo mode, just as in the galaxy population (Bower et al. 2006). The SF and cooling processes in a hierarchical cosmology can account consistently for the different evolution with redshift of the differing luminosity populations.

A second notable feature in Fig. 18 is the strong reduction of the visible as opposed to the total AGN populations. The effect is stronger in the faintest populations where we see a decrease in the space density of AGNs, sometimes by more than an order of magnitude. This behaviour is ascribed to the fact that the faintest populations are subject to very strong obscuration. Thus, their abundances peak at lower redshifts beyond which they decline faster. The strong obscuration also contributes to the flattening of the slope, rendering the evolution of these AGNs quite modest. From an observational point of view, our results suggest that samples of AGNs that are subject to appreciable obscuration can lead to misleading conclusions. For example, in both the Croom et al. (2009b) QSO[4] and the Hasinger et al. (2005) SX AGN samples, no correction for obscuration is adopted. Given that at low redshifts only unabsorbed AGNs are detected in SXs (La Franca et al. 2005), these samples can only trace the evolution of a small fraction of the total AGN population.

When accounting for these effects, our model suggests that the cosmic interplay between the starburst and hot-halo mode results in a complex evolution pattern for the AGN. In this context, low-luminosity AGNs avoid the cosmic fate of their bright counterparts (namely becoming dramatically less numerous with time), because accretion during the hot-halo mode provides gas to enable modest AGN activity in the low-z universe. Note that, as already implied by our analysis, modest activity implies low accretion rates rather than low BH masses. The AGNs in the hot-halo mode show a wide range of BH masses, and therefore in our model, the AGN downsizing *does not* imply that the growth of low-mass BHs is deferred to low redshifts. Average-size BHs accreting at low rates as an interpretation of the AGN downsizing is supported by observations of X-ray-selected AGNs in the *Chandra Deep Field*-South at $z < 1$ (Babić et al. 2007).

Finally, we show in Fig. 19 the density evolution of accretion rate and BH mass for different BH mass ranges (top and bottom panels, respectively). As expected from the discussion in the previous sections, the different accretion channels show different evolution, due to the interplay between the physical mechanisms that govern the two modes. For example, the BH population that grows primarily in the starburst mode varies in mass with redshift and tends to include the most massive BHs[5] (except for the extreme $M_{\rm BH} > 10^9\,{\rm M}_\odot$ BHs). In contrast, the hot-halo mode is always more important for BHs with masses 10^8–$10^9\,{\rm M}_\odot$.

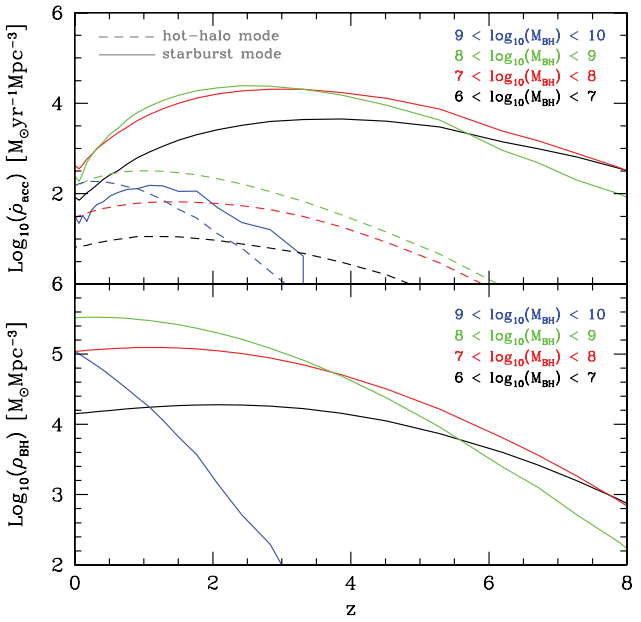

Figure 19. The evolution of the accretion rate (top panel) and BH mass (bottom panel) per unit volume for different BH mass populations. The different colour lines indicate different BH mass populations. In the top panel, we have distinguished between the starburst (solid lines) and hot-halo mode (dashed lines).

The hierarchical growth of BHs is evident in the evolution of the different BH populations. At high redshift ($z > 4$), most of the accreted baryons are deposited in low-mass BHs ($M_{\rm BH} < 10^8\,{\rm M}_\odot$). Thus, the BH mass density is dominated by $10^6 < M_{\rm BH} < 10^8\,{\rm M}_\odot$ BHs. Below $z = 4$, more massive BHs ($M_{\rm BH} > 10^8\,{\rm M}_\odot$) become increasingly numerous as accretion becomes more intense. As a result, the BH mass density at $z = 0$ is dominated by $10^8 < M_{\rm BH} < 10^9\,{\rm M}_\odot$ BHs. Interestingly, according to the bottom panel of Fig. 19, today BHs with masses $10^7 < M_{\rm BH} < 10^8\,{\rm M}_\odot$ and $M_{\rm BH} > 10^9\,{\rm M}_\odot$ have the same mass density.

6 CONCLUSIONS

We have predicted the evolution of AGNs using an extended version of the galaxy formation model GALFORM. Using semi-analytic techniques, GALFORM simulates the formation and evolution of galaxies in a ΛCDM cosmology. In Fanidakis et al. (2011), we developed a physical model of the growth of BHs (including a calculation of spin) and the associated AGN activity that is coupled to the computation of the growth of galaxies and structure in the DM. Here, we have developed and refined this model. The only change made to

[4] These are only type-1 AGNs (Croom et al. 2009a).

[5] The population of BHs with masses 10^6–$10^7\,{\rm M}_\odot$ dominates the accretion rate density (and also BH mass density) at $z > 8$. However, we do not show the predictions of the model at such high redshifts because the resolution of the simulation drops rapidly.

© 2011 The Authors, MNRAS **419,** 2797–2820
Monthly Notices of the Royal Astronomical Society © 2011 RAS

the galaxy formation part of the calculation is to adopt the improved SF law implemented in GALFORM by Lagos et al. (2011c).

We calculate the cosmic evolution of the fundamental parameters that describe BHs: the BH mass, $M_{\rm BH}$, spin, a, and the accretion rate on to the BH, $\dot{m}$ (expressed in units of the Eddington accretion rate). We find that at high redshift ($z \sim 6$), it is mainly the 10^6–$10^7\,{\rm M}_\odot$ BHs accreting at $\dot{m} \simeq 0.3$ that are growing actively. This picture changes at lower redshifts where we find that the accretion activity peaks for 10^7–$10^8\,{\rm M}_\odot$ BHs, accreting at $\dot{m} \simeq 0.05$. Throughout the evolution of these BHs, their spin remains low because of the chaotic fashion in which gas is accreted. However, when these BHs grow to masses of $\gtrsim 5 \times 10^8\,{\rm M}_\odot$, they acquire high spins as a consequence of the burgeoning dominance of mergers with other BHs.

Knowledge of the values of $M_{\rm BH}$, a and $\dot{m}$ allows us to calculate the luminosity produced during the accretion of gas. To do so, we assume that accretion takes place in two distinct regimes: thin-disc (radiatively efficient) and ADAF (radiatively inefficient). Using the luminosities predicted for each AGN (formed when the central BH experiences an accretion episode) and an empirical prescription for the effects of obscuration (Hasinger 2008), we can calculate the LF of all accreting objects in the optical, SXs and HXs. We find good agreement with observations, particularly in the optical, by adjusting the value of the proportionality factor that determines the accretion time-scale in our model.

One of the key predictions of the model is that the brightest AGNs ($\gtrsim 10^{46}\,{\rm erg\,s^{-1}}$) are preferentially found in DM haloes with masses $\sim 10^{12}$–$10^{13}\,{\rm M}_\odot$ in the low-redshift universe ($z \lesssim 2$). This runs contrary to the commonly held view that the brightest AGNs should be hosted by the most massive DM haloes. These intermediate-mass haloes are typical environments where gas-rich galaxies undergo frequent disc instabilities, and occasionally mergers with other galaxies, that provide gas for intense BH growth (starburst mode). More massive DM haloes are usually in quasi-hydrostatic equilibrium and therefore subject to feedback from the AGN. As a consequence, there is less gas in such haloes and typically lower accretion rates, which result in very faint AGN luminosities (hot-halo mode).

The presence of two distinct accretion channels naturally causes a downsizing in the predicted AGN population, which is accentuated on applying obscuration. These channels shape the evolution of the faint AGNs with cosmic time. The faint end of the LF at low z is dominated by massive BHs experiencing quiescent accretion. In contrast, the bright end is always populated by AGNs radiating at near or greater than the Eddington limit. In addition, our model suggests that a significant fraction of AGNs are obscured in the optical and SXs. We find that ∼90 per cent of the total number of AGNs in the $z = 0.015$–4.8 universe are not visible in SXs. The implications of the obscuration are further revealed when we study the cosmological evolution of populations of AGNs of different luminosity. As demonstrated by Fig. 18, low-luminosity AGN populations are strongly attenuated by obscuration. As a result, the peak of AGN activity appears shifted to lower redshifts.

Our analysis explicitly demonstrates that hierarchical cosmological models of galaxy formation and evolution provide a robust framework in which the evolution of AGNs can be studied. The ability of our galaxy formation model to reproduce the observed AGN LFs and account for the mechanisms that shape the evolution of the faint and bright AGN populations extends the capabilities of semi-analytic modelling and shows that the level of AGN activity associated with AGN feedback is compatible with observations. Furthermore, the model makes a number of predictions that can, in principle, be tested by further observations, such as the BH masses and accretion rates associated with different levels of AGN activity or the high fraction of obscured AGNs as a function of redshift. In future analyses, we will explore the spatial clustering and environmental dependences of AGNs as further means to test our model.

ACKNOWLEDGMENTS

NF acknowledges receipt of a fellowship funded by the European Commission's Framework Programme 6, through the Marie Curie Early Stage Training Project MEST-CT-2005-021074. AJB acknowledges the support of the Gordon & Betty Moore Foundation. SC acknowledges the support of a Leverhulme research fellowship. CSF acknowledges a Royal Society Wolfson Research Merit Award. This work was supported in part by an STFC Rolling Grant to the Institute for Computational Cosmology.

REFERENCES

Abramowicz M. A., Chen X., Kato S., Lasota J.-P., Regev O., 1995, ApJ, 438, L37
Adelman-McCarthy J. K. et al., 2006, ApJS, 162, 38
Aird J. et al., 2010, MNRAS, 401, 2531
Alexander D. M. et al., 2003, AJ, 126, 539
Alexander D. M., Chartas G., Bauer F. E., Brandt W. N., Simpson C., Vignali C., 2005, MNRAS, 357, L16
Alexander D. M. et al., 2008, ApJ, 687, 835
Antonucci R., 1993, ARA&A, 31, 473
Babić A., Miller L., Jarvis M. J., Turner T. J., Alexander D. M., Croom S. M., 2007, A&A, 474, 755
Baker J. G., McWilliams S. T., van Meter J. R., Centrella J., Choi D. I., Kelly B. J., Koppitz M., 2007, Phys. Rev. D, 75, 124024
Ballantyne D. R., Shi Y., Rieke G. H., Donley J. L., Papovich C., Rigby J. R., 2006a, ApJ, 653, 1070
Ballantyne D. R., Everett J. E., Murray N., 2006b, ApJ, 639, 740
Barcons X. et al., 2007, A&A, 476, 1191
Barger A. J., Cowie L. L., Mushotzky R. F., Yang Y., Wang W.-H., Steffen A. T., Capak P., 2005, AJ, 129, 578
Baugh C. M., 2006, Rep. Prog. Phys., 69, 3101
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Berti E., Volonteri M., 2008, ApJ, 684, 822
Blandford R. D., Payne D. G., 1982, MNRAS, 199, 883
Blandford R. D., Znajek R. L., 1977, MNRAS, 179, 433
Blitz L., Rosolowsky E., 2006, ApJ, 650, 933 (BR06)
Bongiorno A. et al., 2007, A&A, 472, 443
Bonoli S., Marulli F., Springel V., White S. D. M., Branchini E., Moscardini L., 2009, MNRAS, 396, 423
Bonoli S., Shankar F., White S. D. M., Springel V., Wyithe J. S. B., 2010, MNRAS, 404, 399
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Bower R. G., McCarthy I. G., Benson A. J., 2008, MNRAS, 390, 1399
Boyle B. J., Shanks T., Peterson B. A., 1988, MNRAS, 235, 935
Boyle B. J., Shanks T., Croom S. M., Smith R. J., Miller L., Loaring N., Heymans C., 2000, MNRAS, 317, 1014
Brandt W. N., Hasinger G., 2005, ARA&A, 43, 827
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Cole S. et al., 2001, MNRAS, 326, 255
Comastri A., 2004, in Barger A. J., ed., Supermassive Black Holes in the Distant Universe. Kluwer, Dordrecht, p. 245
Comastri A., Setti G., Zamorani G., Hasinger G., 1995, A&A, 296, 1
Cowie L. L., Barger A. J., Bautz M. W., Brandt W. N., Garmire G. P., 2003, ApJ, 584, L57
Croom S. M., Smith R. J., Boyle B. J., Shanks T., Loaring N. S., Miller L., Lewis I. J., 2001, MNRAS, 322, L29

Croom S. M., Smith R. J., Boyle B. J., Shanks T., Miller L., Outram P. J., Loaring N. S., 2004, MNRAS, 349, 1397
Croom S. M. et al., 2009a, MNRAS, 392, 19
Croom S. M. et al., 2009b, MNRAS, 399, 1755
Croton D. J., 2009, MNRAS, 394, 1109
Croton D. J. et al., 2006, MNRAS, 365, 11
da Ângela J. et al., 2008, MNRAS, 383, 565
Dalla Vecchia C., Bower R. G., Theuns T., Balogh M. L., Mazzotta P., Frenk C. S., 2004, MNRAS, 355, 995
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
Di Matteo T., Springel V., Hernquist L., 2005, Nat, 433, 604
Done C., Gierliński M., Kubota A., 2007, A&AR, 15, 1
Efstathiou G., Lake G., Negroponte J., 1982, MNRAS, 199, 1069
Fan X. et al., 2001a, AJ, 121, 54
Fan X. et al., 2001b, AJ, 122, 2833
Fan X. et al., 2003, AJ, 125, 1649
Fanidakis N., Baugh C. M., Benson A. J., Bower R. G., Cole S., Done C., Frenk C. S., 2011, MNRAS, 410, 53
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Fiore F. et al., 2003, A&A, 409, 79
Fontanot F., Cristiani S., Monaco P., Nonino M., Vanzella E., Brandt W. N., Grazian A., Mao J., 2007, A&A, 461, 39
Fukugita M., Peebles P. J. E., 2004, ApJ, 616, 643
Gebhardt K. et al., 2000, ApJ, 539, L13
George I. M., Turner T. J., Netzer H., Nandra K., Mushotzky R. F., Yaqoob T., 1998, ApJS, 114, 73
Ghisellini G. et al., 2010, MNRAS, 405, 387
Giacconi R. et al., 2002, ApJS, 139, 369
Gilli R., Risaliti G., Salvati M., 1999, A&A, 347, 424
Gilli R., Comastri A., Hasinger G., 2007, A&A, 463, 79
Goulding A. D., Alexander D. M., 2009, MNRAS, 398, 1165
Goulding A., Alexander D., Mullaney J., Gelbord J., Hickox R., Ward M., Watson M., 2011, MNRAS, 411, 1231
Graham A. W., Driver S. P., 2007, MNRAS, 380, L15
Graham A. W., Driver S. P., Allen P. D., Liske J., 2007, MNRAS, 378, 198
Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., 2004, ApJ, 600, 580
Greene J. E., Ho L. C., 2007, ApJ, 667, 131
Greene J. E., Ho L. C., 2009, ApJ, 704, 1743
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Hasinger G., 2004, Nucl. Phys. B, 132, 86
Hasinger G., 2008, A&A, 490, 905
Hasinger G. et al., 2001, A&A, 365, L45
Hasinger G., Miyaji T., Schmidt M., 2005, A&A, 441, 417
Hawley J. F., Gammie C. F., Balbus S. A., 1995, ApJ, 440, 742
Heckman T. M., Kauffmann G., Brinchmann J., Charlot S., Tremonti C., White S. D. M., 2004, ApJ, 613, 109
Hewett P. C., Foltz C. B., Chaffee F. H., 1993, ApJ, 406, L43
Hopkins P. F., Hernquist L., Martini P., Cox T. J., Robertson B., Di Matteo T., Springel V., 2005a, ApJ, 625, L71
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2005b, ApJ, 630, 716
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2005c, ApJ, 632, 81
Hopkins P. F., Hernquist L., Cox T. J., Robertson B., Springel V., 2006, ApJS, 163, 50
Hopkins P. F., Richards G. T., Hernquist L., 2007, ApJ, 654, 731
Ichimaru S., 1977, ApJ, 214, 840
Kauffmann G., Haehnelt M., 2000, MNRAS, 311, 576
Kauffmann G., Heckman T. M., 2009, MNRAS, 397, 135
Kelly B. C., Vestergaard M., Fan X., Hopkins P., Hernquist L., Siemiginowska A., 2010, ApJ, 719, 1315
King A. R., Lubow S. H., Ogilvie G. I., Pringle J. E., 2005, MNRAS, 363, 49
King A. R., Pringle J. E., Hofmann J. A., 2008, MNRAS, 385, 1621
Kochanek C. S. et al., 2001, ApJ, 560, 566
La Franca F. et al., 2002, ApJ, 570, 100
La Franca F. et al., 2005, ApJ, 635, 864
Lagos C. d. P., Cora S. A., Padilla N. D., 2008, MNRAS, 388, 587
Lagos C. d. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011a, preprint (arXiv:1105.2294)
Lagos C. d. P., Padilla N. D., Strauss M. A., Cora S. A., Hao L., 2011b, MNRAS, 414, 2148
Lagos C. d. P., Lacey C. G., Baugh C. M., Bower R. G., Benson A. J., 2011c, MNRAS, 416, 1566
Lapi A., Shankar F., Mao J., Granato G. L., Silva L., De Zotti G., Danese L., 2006, ApJ, 650, 42
Libeskind N. I., Cole S., Frenk C. S., Helly J. C., 2006, MNRAS, 368, 1381
McCarthy I. G. et al., 2010, MNRAS, 406, 822
McLure R. J., Dunlop J. S., 2002, MNRAS, 331, 795
Magorrian J. et al., 1998, AJ, 115, 2285
Mahadevan R., 1997, ApJ, 477, 585
Malbon R. K., Baugh C. M., Frenk C. S., Lacey C. G., 2007, MNRAS, 382, 1394
Marchesini D., Celotti A., Ferrarese L., 2004, MNRAS, 351, 733
Marconi A., Hunt L. K., 2003, ApJ, 589, L21
Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, MNRAS, 351, 169
Martínez-Sansigre A., Rawlings S., Lacy M., Fadda D., Marleau F. R., Simpson C., Willott C. J., Jarvis M. J., 2005, Nat, 436, 666
Marulli F., Bonoli S., Branchini E., Moscardini L., Springel V., 2008, MNRAS, 385, 1846
Menou K., Haiman Z., 2004, ApJ, 615, 130
Merritt D., Milosavljević M., Favata M., Hughes S. A., Holz D. E., 2004, ApJ, 607, L9
Miyaji T., Hasinger G., Schmidt M., 2000, A&A, 353, 25
Monaco P., Fontanot F., Taffoni G., 2007, MNRAS, 375, 1189
Narayan R., Yi I., 1994, ApJ, 428, L13
Norberg P. et al., 2002, MNRAS, 336, 907
Parkinson H., Cole S., Helly J., 2008, MNRAS, 383, 557
Parry O. H., Eke V. R., Frenk C. S., 2009, MNRAS, 396, 197
Piconcelli E., Guainazzi M., Cappi M., Jimenez-Bailon E., Schartel N., 2005, A&A, 432, 835
Plotkin R. M. et al., 2010, AJ, 139, 390
Rees M. J., 1982, in Riegler G. R., Blandford R. D., eds, AIP Conf. Ser. Vol. 83, The Galactic Center. Am. Inst. Phys., New York, p. 166
Reynolds C. S., 1997, MNRAS, 286, 513
Richards G. T. et al., 2005, MNRAS, 360, 839
Richards G. T. et al., 2006, AJ, 131, 2766
Risaliti G., Maiolino R., Salvati M., 1999, ApJ, 522, 157
Ross N. P. et al., 2009, ApJ, 697, 1634
Schmidt M., Green R. F., 1983, ApJ, 269, 352
Schmidt M., Schneider D. P., Gunn J. E., 1995, AJ, 110, 68
Schulze A., Wisotzki L., 2010, A&A, 516, A87
Shakura N. I., Sunyaev R. A., 1973, A&A, 24, 337
Shankar F., 2009, New Astron. Rev., 53, 57
Shankar F., Salucci P., Granato G. L., De Zotti G., Danese L., 2004, MNRAS, 354, 1020
Shankar F., Weinberg D. H., Miralda Escudé J., 2009, ApJ, 690, 20
Shen Y. et al., 2009, ApJ, 697, 1656
Simpson C., 2005, MNRAS, 360, 565
Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481
Springel V. et al., 2005a, Nat, 435, 629
Springel V., Di Matteo T., Hernquist L., 2005b, MNRAS, 361, 776
Steffen A. T., Barger A. J., Cowie L. L., Mushotzky R. F., Yang Y., 2003, ApJ, 596, L23
Thacker R. J., Scannapieco E., Couchman H. M. P., 2006, ApJ, 653, 86
Treister E., Urry C. M., 2006, ApJ, 652, L79
Tremaine S. et al., 2002, ApJ, 574, 740
Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886
Urry C. M., Padovani P., 1995, PASP, 107, 803
Vasudevan R. V., Fabian A. C., 2007, MNRAS, 381, 1235
Vestergaard M., 2002, ApJ, 571, 733
Vestergaard M., Osmer P. S., 2009, ApJ, 699, 800

Vestergaard M., Peterson B. M., 2006, ApJ, 641, 689
Vestergaard M., Fan X., Tremonti C. A., Osmer P. S., Richards G. T., 2008, ApJ, 674, L1
Volonteri M., Sikora M., Lasota J.-P., 2007, ApJ, 667, 704
Warren S. J., Hewett P. C., Osmer P. S., 1994, ApJ, 421, 412
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Willott C. J. et al., 2010, AJ, 140, 546
Wisotzki L., Christlieb N., Bade N., Beckmann V., Köhler T., Vanelle C., Reimers D., 2000, A&A, 358, 77
Wolf C., Wisotzki L., Borch A., Dye S., Kleinheinrich M., Meisenheimer K., 2003, A&A, 408, 499
Yu Q., Lu Y., 2008, ApJ, 689, 732
Yu Q., Tremaine S., 2002, MNRAS, 335, 965

This paper has been typeset from a TEX/LATEX file prepared by the author.